Deliverable 4.1 Specification of user profiling and contextualisation

Dorothea Tsatsou, Maria Loli, Vasileios Mezaris (CERTH)

Rüdiger Klein, Manuel Kober (FRAUNHOFER)

Tomáš Kliegr, Jaroslav Kuchar (UEP)

Matei Mancas, Julien Leroy (UMONS)

Lyndon Nixon (STI)

18.4.2012

Work Package 4: Contextualisation and Personalisation

LinkedTV

Television Linked To The Web

Integrated Project (IP)

Dissemination level[1]	PU
Contractual date of delivery	31th March 2011
Actual date of delivery	18h April 2011
Deliverable number	D4.1
Deliverable name	*Specification of user profiling and contextualisation*
File	*LinkedTV_D4.1_Specification of user profiling and contextualisation.doc*
Nature	*Report*
Status & version	*Final*
Number of pages	*84*
WP contributing to the deliverable	*WP 4*
Task responsible	*CERTH*
Other contributors	*Fraunhofer* *UEP* *UMONS* *STI*
Author(s)	*Dorothea Tsatsou, Maria Loli, Vasileios Mezaris* *CERTH* *Rüdiger Klein, Manuel Kober* *Fraunhofer* *Tomáš Kliegr, Jaroslav Kuchar* *UEP* *Matei Mancas, Julien Leroy* *UMONS* *Lyndon Nixon* *STI*

---

[1] • PU = Public

• PP = Restricted to other programme participants (including the Commission Services)

• RE = Restricted to a group specified by the consortium (including the Commission Services)

• CO = Confidential, only for members of the consortium (including the Commission Services))

Reviewer	*Lyndon Nixon, STI*
EC Project Officer	*Manuel Carvalhosa*
Keywords	*semantic, ontologies, user profile, user model, personalisation, contextualisation, behavioural tracking*
Abstract (for dissemination)	*This deliverable presents a comprehensive research of past work in the field of capturing and interpreting user preferences and context and an overview of relevant digital media-specific techniques, aiming to provide insights and ideas for innovative context-aware user preference learning and to justify the user modelling strategies considered within LinkedTV's WP4. Based on this research and a study over the specific technical and conceptual requirements of LinkedTV, a prototypical design for profiling and contextualizing user needs in a linked media environment is specified.*

Table of contents

List of Figures

List of Tables

List of Abbreviations

CF	Collaborative Filtering
CUM	Contextualised User Model
UTA	UTilités Additive
LOD	Linked Open Data
DL	Description Logic
CSME	Condat's Smart Media Engine
WUM	Web Usage Mining
POS	Part-of-Speech
SVM	Support Vector Machines
BOW	Bag-of-Words

1 Introduction

In order to manage the huge amount of information available today people need targeted support by intelligent information systems. This issue arises in many fields: print media, electronic media, digital media, the WWW, marketing and advertising, eLearning, tutoring systems, etc. Targeting information is dependant on the different needs, preferences and situations of the system's users. Consequently, one of the key challenges in intelligent information systems is personalisation and contextualisation. The process of personalizing and contextualizing the information pertinent to the user in such a system is divided to two distinct aspects: a) capturing and representing user preferences and context and b) using this knowledge about the user to filter the diverse information provided by the system and present the most interesting information given the distinct user situation.

The scope of this deliverable is oriented towards addressing the first of these two aspects that will subsequently guide the actions in future work on the second aspect. In general and independently of the focus application, in order to interpret and adapt the information provided by user-pertinent features the system typically employs a user model. Overall, a user model is an information structure characterizing the user in such a way that allows the system to find out if a given "piece of information"/content item – e.g. a document, a video, a scene, an article etc. – could be interesting for him or not.

The user model is neither static nor just a mere accumulation of preferences, hence user preference learning, i.e. understanding, composing and continuously adapting the user model to new, complex information about the user and his behavior, is of pivotal importance for personalisation. An intelligent user modeling system should be able to adapt to additional information provided by the user explicitly and/or implicit information derived through observations of user behaviour.

In fact, as we will argue in chapter 2, current personalisation approaches are oriented towards unobtrusively extracting and understanding implicit user feedback and translating it in a machine-understandable format appropriate for making predictive inference about relevant content. Preference learning may take into account the user's transactional history, information derived from his social networks activity and/or available semantic background knowledge.

Moreover, the preferences of a given user evolve over time and various states of the user model are applicable in different circumstances. Therefore, the context in which the user is at a given moment has a significant impact in efficient decision-making in personalised environments. Context refers to the certain circumstance under which a user is at a given moment, which influences the relevance of recommended content and can have different aspects: the concrete time (of the day, of the year); the location (at home, at work, on holidays, on business travel); seasonal events (e.g. presidential elections, football championship, a movie just started in the cinemas); activities in the user's social network; the user's reaction to the content and other relevant stimuli.

The question is how this context information adapts to the more stable user model. We call this "merge" the *contextualised user model* (CUM). Obviously, context will influence the user's interests in various domains or concrete entities. We will investigate appropriate means to generate the *contextualised user model* by combining the user model and the user's context.

In networked media environments such as LinkedTV in particular, where media convergence becomes more and more a reality, content providers are impelled to uniformly interconnect the content for consumer usage, whether that might be text, audio, video or interactive Websites, without mediation, while tailoring it to the user's needs. Consequently, an additional challenge arises in capturing user preferences and context: intelligent interpretation of the vastly heterogeneous information encompassed in digital media and sophisticated representation in such a manner that can both coalesce the diverse user-pertinent information under a uniform vocabulary and render them usable for subsequent matching to the disparate linked media content.
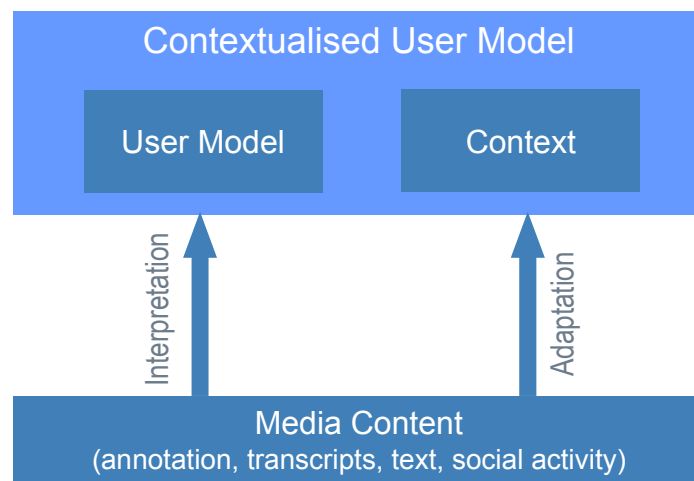
Figure 1: High-level conceptualisation of a contextualised user model

In this deliverable we will examine and address methodologies for personalizing and contextualizing user preferences, mainly focusing on identifying and handling the key requirements within the LinkedTV linked media environment and specifying a prototypical work plan towards achieving identified personalisation and contextualisation goals. More specifically we will:

- Explore the state of the art in user profiling and preference learning and identify the most effective approaches to be extended in LinkedTV (Chapter 2).
- Explore the state of the art in behavioural tracking and contextualising user preferences and identify the most effective approaches to be extended in LinkedTV (Chapter 3).
- Determine the requirements posed within the LinkedTV platform on user profiling and specify proposed approaches to meet with these requirements, either by extending related work or by developing novel implementations based on (Chapter 4).

- Determine the requirements posed within the LinkedTV platform on semantically analysing multimedia content and specify proposed approaches to meet these requirements (Chapter 4).
- Determine the requirements posed within the LinkedTV platform on contextualising user reactional and transactional behaviour and specify proposed approaches to meet these requirements (Chapter 5).

1.1 History of the document

Table 1: History of the document

Date	Version	Name	Comment
2012/1/27	V0.0	Dorothea Tsatsou	First version, merging Fraunhofer's position paper to the D4.1 deliverable structure
2012/3/3	V0.1	Dorothea Tsatsou	First complete draft for WP4 partner review, including contributions from Fraunhofer, UEP, UMONS and STI.
2012/3/15	V0.2	Dorothea Tsatsou	Refined draft after Fraunhofer review
2012/3/19	V0.3	Dorothea Tsatsou	Refined draft after UEP review
2012/3/20	V0.4	Dorothea Tsatsou	Refined draft for QA
2012/4/5	V0.5	Dorothea Tsatsou	Refined final version for QA
2012/4/13	V1.0	Dorothea Tsatsou	Final version after QA and partners' final input.

1.2 Objectives

This deliverable aims to provide **insights** for innovatively profiling and contextualising user preferences, with a focus on networked media environments, based on the past and most current work in the field of personalisation, as well as to make concrete decisions over the **requirements** posed within the LinkedTV scope, **specify** potential personalisation approaches and **justify** proposed decisions for profiling and contextualisation within the LinkedTV platform based on past work. More specifically the deliverable will address related work, requirements and specifications for determining:

- the appropriate information sources for extracting representative user preferences;
- requirements and methodologies for understanding and expressing user preferences in a machine-understandable format;
- appropriate methodologies for updating, learning and maintaining representative user profiles

---

- optimal vocabularies and schemata for representing user preferences;
- methodologies for extracting different contextual situations of the user;
- methodologies for understanding and learning user context;
- appropriate methods to determine contextualised user behaviour;
- optimal schemata for representing contextualised user preferences;
- storage and server communication strategies of the contextualised user profile to ensure privacy safeguarding and profile usage efficiency.

To this end, this deliverable will explore the state of the art in order to pinpoint the most prominent information sources from which preferences and context facets representative to a user accrue and identify salient preference learning techniques pertinent to the LinkedTV scope. In addition, it will examine opportunities and intricacies pertaining user preference extraction and learning in the field of digital media and argue that semantic user profiling and contextualisation is the most suitable approach to cope with the variety of information in networked media environments. Consequently it will address the main facets for inducing a semantic contextualised profile by:

- Determining the granularity and expressivity level of required ontological knowledge to better capture user preference semantics.
- Providing an overview of the design principles for the ontological knowledge most suitable to be used in LinkedTV to capture semantic user preferences.
- Discussing methods to unobtrusively extract user behaviour information in the LinkedTV platform,
  - based on interaction with the content;
  - based on reaction to the content.
- Discussing methods for understanding user preferences and their application context, including:
  - Semantically classifying user preferences, i.e. determining how to convey knowledge about the user in a uniform, machine-understandable (ontological) format.
  - Semantically classifying user context, i.e. determining how user context affects user preferences
- Providing an overview of methodologies for dynamically learning user preferences.
- Providing an overview of methodologies for dynamically learning user context.
- Formally representing contextualised user preferences in a semantic, machine-understandable format, suitable for semantic inferencing.

1.3 Motivation: contextualised user model use cases in digital TV

While refraining from deeply analysing the target usage of CUMs in LinkedTV, as it is out of the scope of the current deliverable and pertinent to future work, we must underline the motivating use cases according to the requirements of which the LinkedTV CUM will be designed.

In general, in a networked media environment, specifically digital television, there are several use cases where a contextualised user model might be applied, some of the most prominent being:

- Content recommendation: While a user is watching a video, he may want to receive additional information related to the content item. He may interrupt the current video and navigate through various links, following recommendations provided by the system. Alternatively, he may bookmark content items (such as specific media fragments or the entire video) and follow related information after the video has been finished based on the concepts describing the video. In this case, the user model will guide the system to recommend related information. It can subsequently highlight or directly deliver relevant videos or textual content, such as trivia, news, events, reviews, advertisements, etc, according to the user's preferences. Furthermore, it might determine the degree of relevance of recommended content to the user by ranking recommendations.

- User adapted versions: A video (e.g. a documentary) may deploy its story using different media segments corresponding to different user interests, skills, priorities, etc and may be configured according to a concrete user's preference model. A similar approach may also work for entertainment and fiction videos where a user may prefer a certain version of the video out of a whole family of pre-defined versions.

- Media search: Today's search engines with their mainly string-oriented search like Google use forms of personalisation for prioritizing search results. They use information from previous search queries (just recent ones or those submitted over a longer time) and infer to some degree the user's interests and preferences.

While the contextualised profiling schema considered in LinkedTV will be implemented at a conceptual level, thus rendering it adaptable to all three of the aforementioned use cases, the objectives of LinkedTV are not oriented towards media search or changing the user experience of browsing to the actual TV programme/video. Since the project's scenario is focused towards linking media content such as full videos and media fragments with other relevant media content and miscellaneous extra web content based on their semantic (concept) characterization, the first use case is highlighted as the target focus for the CUM. To this end, the CUM is aimed to be used by suitable filtering mechanisms that will enable sorting out the most user-pertinent concepts and identifying potentially interesting content items (either in-platform media or external) for the user based on their semantic description.

2   State of the art in user modelling

The user model is a structure containing information about the user, in a uniform, machine-understandable format, employed by the system to support understanding and management of this information in order to enable intelligent content delivery and provide comprehensive recommendations to the user. For this purpose, a user model may contain a characterization of the user's interests and/or disinterests, based on biographic and demographic data, knowledge and skills, interaction with the content, interaction with other users etc. User modeling is the process of creating, learning and maintaining the user model.

In general *"a user model is the knowledge about a user that a system stores which contains explicit assumptions on all aspects of a user that might be relevant to the conduct of the dialogue system and its representation. Assumptions of a system are the objectives of the user (within the scope of the system), the plans with which the user wants to achieve its goals and the knowledge or beliefs (expertise level) of the user within the scope"* [Kobsa85].

This section will focus on exploring the state of the art on the most prominent techniques in modeling user information, initiating from the general aspects of user profiling applicable to any application domain, such as the examination of the sources of information potent to characterise a user and the most salient approaches to take advantage of this information in order to personalise a user's experience in a targeted information delivery platform (chapter 2.1). Further on, we will elaborate on specific strategies suitable for learning complex knowledge about a user in any data space (chapter 2.2). After covering the general principles in the field of preference extraction and learning, we will highlight particular approaches in profiling users in digital media applications in order to identify the most specific challenges pertinent to LinkedTV (chapter 2.3). In chapter 2.4 we will focus on methodologies for implicitly extracting and maintaining a semantic user profile while arguing that this paradigm is the most suitable approach for unobtrusive and meaningful learning of representative user models in a domain of heterogeneous data such as the LinkedTV networked media environment.

2.1   Profiling methodologies

The information to be managed in the user profile tends to be quite complex. The user model captures the user's likes and dislikes in a given domain or a cross-domain environment, while identifying the perceptions and intentions of the user. Four types of resources may be used to determine the user's preferences: information explicitly provided to the system, content features, similarities with users in a social networking environment and explicit perceptions of the world based on predefined knowledge.

2.1.1   Profiling explicit preferences

The most straightforward way to determine user preferences is via explicit information that a user might provide upon registration to a web site/service [Srivastava00]. Such information

would include demographic details (e.g. age, sex, family status, ethnicity, place of birth current location), information on the user's skills and credentials (e.g. education, work details), personal attributes (e.g. hobbies, political and religious views) or a user-defined selection of available content and/or categories. Explicit information provides a strong and concrete understanding about the user that could aid subsequent content filtering.

As an example, [Pannu11] has proven the efficiency in web search filtering while employing an explicit user profile in contrast to traditional search. Their mediation system uses vector space models (VSM), where a vector space is created, stored and update for a user, containing explicit user interests and any additional personal information. In [Tzviskou07] the system combines declared identification information and registered user requests of ontological objects during his navigation in a web platform to optimize web navigation.

However, while most social networks and other personalised services nowadays allow and prompt users to provide explicit information about themselves, early studies have shown that in general users are *reluctant to disclose sensitive personal information* on the web, especially in terms of their demographic data (age, profession etc) [Bae03], [Kobsa01], or even *indisposed to put in the effort* to provide preference-specific information. As a consequence to the unwillingness of users to provide explicit feedback, current personalisation strategies do not expect users to explicitly state their preferences [Mulvenna00] but are rather oriented towards implicit preference mining and learning. Nevertheless, implicitly capturing and managing user information, mostly by means of tracking their transactions on an information system, gives rise to different privacy compromising issues that an intelligent implicit personalisation system must address.

2.1.2 Content-based profiling

Content-based preference learning takes into account features of the available content in combination with the history of a user's transactions with these features [Pazzani07]. Typically, content-based profiling systems consists of two components: the content analyzer, where descriptive features are extracted from raw unstructured content and the profile learner which resorts to machine learning techniques in order to generalize the data collected from the user's transaction history and infer a predictive user model [Lops11].

A widely popular implicit content-based learning technique, while however based on explicit feedback, is analyzing the history of ratings that a user might have assigned to consumed content and its relevance with other content items using existing "content descriptions to uncover relations between items in order to later predict unknown ratings" [Amatriain09], i.e. ratings for new content that might be introduced to the user.

Alternatively, obviating from the need for explicit user action, other approaches employ text mining, i.e. analyse the textual attributes (tokenized text, keywords etc) of consumed content and use them to characterise visited documents, while using this characterisation to learn the user profile. For instance, the "Stuff I've seen" system [Dumais03] promotes information re-use by indexing user consumed content based on the linguistic analysis of visited text. The

index provides fast access on information included in the user transaction history while enriches web searches with contextual information.

[Papadopoulos09] use lexical graphs carrying statistical information, such as frequency and co-occurrence between terms in textual content (articles, keywords, ads etc), which are employed to expand content information in order to leverage the *vocabulary impedance* problem in constrained content descriptions and provide a better understanding of the content. This problem was first addressed in the work of [Ribeiro05] to describe vocabulary scarcity in short advertisement descriptions. Arguably, the vocabulary impedance problem is extended to all kinds of content for which descriptions (metadata, tags, text) are sparse and ambiguous and thus affects digital media environments.

The challenges addressed in the aforementioned overview indicate that the effectiveness of implicit content-based preference mining heavily depends on *expressive descriptions of content items* or *advanced textual analysis* in order to extract comprehensive content features, combined with *adequate information about the user* (transaction history).

2.1.3 Peer-based profiling

Another significant source of information about a user is the interaction between peers within the system itself or within a social media environment. Therefore, for some tasks it is valuable to identify typical categories of users that use the system in a similar way, expect from it similar results and can be described by similar sets of features. Similarly to rating-based content approaches, the performance of peer-based preference discovery is highly dependant on the volume of peers of a given user.

Collaborative-based profiling

Collaborative filtering (CF) is the most widely considered technique in the field of peer-based preference mining and recommendation. In CF the behaviour of co-interacting user groups is analyzed to infer the interests of individuals in this collaborative environment. Extracting meaningful information in such an environment predominantly requires very large amounts of peer-retrieved data, namely by collecting information about the behaviour and preferences of as many users as possible.

Collaborative preference learning strategies focus on determining synergy between particular user groups based on collective user behaviour over the content (e.g. content ratings, click-through). A general definition is that *"user preferences are inferred from past consumption patterns or explicit feedback and predictions are computed by analyzing other users"* [Karatzoglou10]. Two main implementations of collaborative filtering (CF) are distinguished: the neighborhood methods and the latent factor models.

The former calculates the similarity between users and recommends the similar users' items. User-based approaches associate a set of nearest neighbors to each user (e.g. via a *Nearest-Neighbor* Algorithm [Resnick94]), and then predicts a user's rating on an item using the ratings for that item of the user's nearest neighbors. A faster model-based [Breese98] variant of collaborative filtering relies on clustering [Resnick94] [Kleinberg04] [Kelleher03]

which involves firstly constructing similar user groups, and then predicting a user's rating on an item by taking into account the ratings for that item assigned by the members of his group.

Latent factor models are solved in a more complicated, algebraic way which involves profiling both the user and the content items and decomposing these profiles in a unified latent factor space [Koren08]. This approach is described in multiple recent works such as [Koren09], [Mairal10] and in a social version in [Ma08].

Regardless the implementation, *data sparcity*, low *prediction quality* and, more importantly, *scalability* were always the most crucial challenges for collaborative filtering [Ma08], the latter leading CF approaches to resort to offline users and/or items clustering [Mobasher07].

Social network mining

Within social networks [Jamali2010] the preference learning system aims to understand if the people that a user is connected to have certain interests that frequently also apply for the seed user. Extracting user information in a social network differs from traditional CF in that the system is already aware of (social) connections between users. A social network mining system attempts to propagate interest within the network by employing both CF and general graph-based data mining techniques. In a large social network with many users and many more connections, it is subsequent that the mining methodologies will be hampered by the same *scalability* issues as in CF. Nevertheless, through social network mining, it is possible to extract significant knowledge about the preferences, demographic data, online activities, needs and expectations of an individual social network user.

 [Yang11] justify that combining interest correlation from user-service interactions and friendship interconnections in social networks provides higher performance in service and relevant users recommendations. Their friendship-interest propagation framework uses both the neighbourhood and latent factor CF techniques to cluster aggregated user interests and a random walk algorithm for homophily discovery produces a unified prediction model but the authors have yet to discover how this model impacts individual users' decision making.

Similarly, [Zhang11] extend uniform influence diffusion within users in a social network by discovering the top influential users for a related user in a specific domain based on aggregated user preferences. These techniques can be used to discover the most relevant preferences for an individual user based on the influence of their social connections.

An approach to detect interest interconnections between users in social networks is to identify the relations between descriptions of the content (keywords, tags, concepts) based on their participation in user communities. This method, known as community detection, employs graph theory methods to identify sets of nodes in folksonomies [VanderWal04], most commonly manifested as tripartite hypergraphs, in which node types describe the users, the tags and the resources [Liu10], that are more densely connected to each other than to the rest of the graph [Newman03]. The correlations between the users, items and resources can be used to infer aggregated interest models within these communities.

Community detection and other tripartite graph-based clustering methods (e.g. clique discovery [Balasundaram11]) in folksonomies are particularly widespread in social tagging systems, popularized by web 2.0: users tag resources on the web (pictures, video, blog posts etc.) The set of tags forms the folksonomy which can be seen as a shared vocabulary that is both originated by, and familiar to, its primary users [Mika07]. Ontologies have also been designed to capture and exploit the activities of social tagging [Gruber05] [Kim07] [Passant07] while researchers have attempted to bridge folksonomies and ontologies to leverage *data sparsity* by discovering *the semantics of tags*.

2.1.4 Knowledge-based profiling

Knowledge-based preference learning techniques aim to leverage the lack of information when new users or new content are introduced in a preference learning system, identified as the *cold-start problem* [Middleton04], by resorting to explicit semantic background knowledge. This knowledge provides compact, uniform and structured domain information, appropriate for intelligent recommendation inferencing services. The use of ontological (or simply taxonomical) knowledge in order to improve recommendation accuracy and completeness has been explored widely in the past.

[Bhowmick10] uses ontology-based profiles for personalised information retrieval, employing both a static facet and a dynamic facet for the profile, allowing them to capitalise on the expressivity and adaptability of formal semantic conceptualisations. However no specific method for implicitly understanding the semantics of raw contextual information is defined. The adaptability and flexibility potentials offered by the uniform descriptions in ontology-based profiling methodologies render them suitable candidates for personalisation in context-specific systems and even in resource-limited mobile environments [Weissenberg04].

In [Kearney05] an "impact" factor is introduced to measure the influence of ontology concepts to an ontological user profile in order to trace user behavioural patterns in web navigation. The retrieved URLs are assumed to be mapped directly to ontology concepts. [Belk10] describes an Ontological Cognitive User Model (OCUM), used to depict the unique cognitive parameters of a user. The goal is to provide accurate representations of user profiles and exploit the potential to semantically map hypermedia content effectively in order to enhance profile adaptation and web service navigation.

Popular knowledge-based preference mining techniques map user interests onto the ontology (or taxonomy) itself or an instance of the ontology by activating concepts and assessing their impact based on the interaction of the users with content described by predefined concepts. In [Trajkova04] and [Sieg07], variants of the spreading activation algorithm are employed to propagate weighted interests up the hierarchy of a taxonomy. These ontology-*based* user profiles however lack the ability to express more complex, axiomatic semantic information about user behavioural patterns and are further memory or server communication dependent since they require storage of the full ontology on the client in order to reflect the user profile.

2.1.5 Hybrids

As aforementioned, the quality of content-based and collaborative based profiling methods depend highly on the *existence of adequate information* about the content items or the user correspondingly. On the other hand, knowledge-based systems rely on the existence of background knowledge that requires experts and *manual work* as well as the *existence of comprehensive semantic mappings* of the content denotation to the ontology. Several profile learning techniques use a hybrid approach by combining the aforementioned implicit preference extraction methods, which helps compensate the limitations of the individual systems [Burke07].

The alliance of content-based and knowledge-based techniques has most notably been explored for the purpose of complementing vague transaction history profiles with intelligent ontological knowledge and to compensate with the lack of semantic mappings between sparsely annotated content and ontologies. For instance in [Sieg07] preference propagation in the ontology-based profiles of users initiates from the concepts emerging from observation of the navigational behaviour of the users by taking into account interest factors such as "the frequency of visits to a page, the amount of time spent on the page and other user actions such as bookmarking".

In [Gemmis08] the authors capitalise on the synergy between user generated content encompassed in folksonomies and provided publisher information over the content to discover potential semantic interpretations of user interests. The semi-structured content information are analysed by machine learning techniques to discover interesting concepts and a probabilistic user model is formed via supervised learning of user-rated content.

[Sieg10] extends standard collaborative filtering preference learning techniques to infer user similarities based on their interest scores across ontology concepts rather than explicit item ratings, thus significantly outperforming traditional CF techniques in prediction accuracy and coverage.

[Tsatsou09] combines content-based and knowledge-based approaches to tackle both the vocabulary impedance and cold-start problems by creating a semantic user profile through observation of the transactional history of the user. The authors perform semantic analysis of domain content by means of a lexical graph, used to automatically interpret and annotate consumed and provided content. A semantic user profile is learned dynamically and updated over time, its axiomatic representation schema allowing for inferencing with standard reasoning engines. This approach was extended in [Tsatsou11], in which the use of community detection for analysing the domain is explored, thus rendering the system able to receive aggregated preference information from social networks.

2.2 Preference learning

Following the identification of user preferences, the goal of preference learning is to induce predictive preference models from empirical data [Fürnkranz10]. The most significant aspect in complex user preference learning is user behaviour pattern discovery. To this end, the

system mines patterns from the data, which represent observed usage regularities and user attributes. There are several groups of machine learning techniques oriented towards addressing this challenge, including clustering, association rules, sequential patterns, and latent variable models. Refraining from an exhaustive reference to the various strategies, this chapter will provide an overview of several generic preference learning approaches that can most suitably exploit the information stemming from the aforementioned preference extraction strategies and are deemed as pertinent within the scope of LinkedTV.

2.2.1 Stereotype models

Social networks provide information about their users including their biographic data, their preferences, and their links to other people. This information can be used to form categories *"constituting strong points of commonalities"* [Kay94] among users called stereotypes. Stereotypes can be used to expand the initial knowledge about the user according to information available about "similar" users. Often, stereotypes form a hierarchy, where more specific stereotypes can inherit some information from their parents, as in the classic system Grundy [Rich79].

A popular way of stereotype-based user modelling is a linear set of categories for representing typical levels of user skills. For example, [Chin89] describes the KNOME system modeling users' expertise of the UNIX operating system on one of four levels (novice, beginner, intermediate, and expert).

Some classic systems, like Grundy [Rich79], have exploited hierarchies of stereotypes long ago. Such structures facilitate stereotype management and make the stereotype transition more effective. The more specific stereotypes inherit characteristics from the more general ones, hence avoiding unnecessary updates. Systems implementing an ontology of stereotypes access the pool of semantic technologies for stereotype processing. [Nebel03] propose the implementation of a database of reusable user ontologies as a source for interoperable user information.

Several systems applying ontologies for stereotype user modelling rely on a domain ontology to populate the characteristics of predefined stereotype profiles. The models of individual users can be represented as overlays based on the same domain ontology. [Gawinecki05] presents an ontology-based travel support system. The system utilizes two ontologies for describing travel-related domains in which particular preferences of a user are inherited from stereotype profiles. The stereotype profiles themselves can be populated based on ontologies. This project signifies the benefit of re-using ontologies – the employed ontologies have not been developed by the authors but reused from relevant projects.

Stereotype-based user modelling is advantageous when from a little hint about a user the system should infer a great deal of modelling information. However, for modelling fine-grained characteristics of individual users (for example, a knowledge level of a particular concept) *more precise overlay models* should be employed.

2.2.2 Object ranking preference learning

The majority of feedback from the user is assumed to be derived from user tracking data rather than conveyed explicitly by the user. While user preferences come from multiple sources (e.g. level of user interaction with the video content, emotions, etc.) they can be aggregated into a single scalar value for the given consumed/viewed content item. The preference order of content-items in user history can be inferred from this scalar number. The input for a preference order learner will consist of a set of objects - tuples consisting of content item feature vector and a user preference level. Building a predictive model from such an input corresponds to *an object ranking preference learning problem*.

There are two prominent approaches to addressing this class of preference learning problems: *learning preference relations* and *learning utility functions*. It is out of the scope of this document to give a comprehensive survey of the state-of-the-art in this area, nevertheless a good general overview of instance and object ranking methods is given in e.g. [Furnkranz10]. We selected two algorithms that we find relevant in the LinkedTV context. As an example of algorithm learning preference relations we selected the PrefLearn algorithm [Dembczyński10]. As a representative of utility learning, we will detail the UTA (UTilités Additives)-class methods [Jacquet82], which learns an additive piece-wise linear utility model. Despite its relatively "old age", UTA method is widely used as a basis for many recent utility-based preference learning algorithms, e.g. [Greco08].

Learning utility functions

In the LinkedTV framework, it is intended to use the explicit but foremost the implicit feedback from an individual user to sort the individual content items that the user has interacted with according to the degree of her interest. Such a representation of user preference – alternatives sorted from the most preferred alternative to the least preferred one – corresponds to the stated order of alternatives, as known from the field of Multicriteria Decision Analysis [Siskos05]. One of the most suitable approaches to evaluating data of this kind is constituted by utility based methods [Furnkranz10]. Of particular interest is the UTA method, which is a disaggregation-aggregation method for preference learning and rule-based methods.

UTA has two important advantages compared to most other preference learning methods: First, the output of the method is a model explaining user preferences in the form of utility functions, which is perhaps the most widely known representation of preferences. Utility functions are covered by elementary college or high-school microeconomic courses. In this respect, UTA method best matches the requirement that the user should be able to inspect her preference model. Utility curves learnt by the UTA method can be easily visualized and prospectively even edited, as for example implemented in the Visual UTA software [Grycza04].

Second, since the UTA method has a very strong inductive bias, the number of required training examples is smaller than for other methods. This is a significant advantage given the fact that the user starts with no training examples.

Many preference learning methods lead to hard optimization problems. The UTA method has the advantage that the learning process can be posed as linear convex program.

Learning preference relations

The *PrefLearn* algorithm is based on the principle of rank loss minimization. The goal of the algorithm is to learn a set of decision rules of the form:

> *if the difference in the strength of "Berlin"* between content items x and x' is >= 0*
>
> *and the difference in length between content items x and x' is more than 1 minutes,*
>
> *then content item x is preferred over content item x' with credibility 0.48.*
>
> *(*"Berlin" being a concept or instance in a reference ontology)*
>
> *(adapted on envisaged LinkedTV input based on: [Dembczyński10])*

In a testing phase, ranking of objects is done by using the NFS (Net Flow Score) procedure, which assigns a score based on the contribution of individual rules matching the object. The objects (in LinkedTV context the content items) are then sorted according to the NFS score.

The advantage of the PrefRules algorithm is that it is very concise while maintaining good performance. The error rate for the PrefRules algorithm is below a state-of-the art SVM classifier is reportedly achieved with a small number of rules. This comes, however, at much higher computational cost [Dembczyński10].

2.2.3 User clustering

While in section 2.1.3 we explored mining preferences for the individual user based on peers' information, this section refers to clustering web user preferences with the purpose of implicitly harvesting collective intelligence in order to identify common usage patterns. These patterns can arguably be used to infer general purpose, top-level conceptual relations and rules over subsets of the domain knowledge.

There is a wide range of literature on clustering web users, a recent survey of the key algorithms can be found in [Liu11]. Here, we exemplify a single algorithm that we consider particularly relevant to LinkedTV, which was introduced within the Lumberjack framework in paper [Chi02]. It is a multi-modal clustering algorithm, and as such it is suitable for incorporating information from multiple sources.

Individual user profiles are assembled by aggregating the feature vectors of the pages in the visitor's clickstream. The collective user profiles are clustered using Repeated Bisectional K-Means with cosine measure as a distance metric. This version of the K-means algorithm starts with putting all user profiles into one initial cluster and then repeatedly bisects clusters using the traditional K-means algorithm until a pre-specified number of clusters is obtained.

As to the choice of cosine similarity as the distance function, it is more common to use Euclidean distance with the K-Means algorithm. However, it was shown in [Strehl00] that Cosine measure, which is widely used in information retrieval, gives much better results for

high dimensional data such as shopping baskets and textual documents than the Euclidean distance.

In the original paper, all the modalities came from the hypertextual environment, but it is easily conceivable that this can be extended to video/audio modalities of multimedia content items with little modifications.

Another approach to detect interest interconnections between users in social networks in particular is to identify the relations between descriptions of the content (keywords, tags, concepts) based on their participation in user communities, i.e. community detection which was further described in chapter 2.1.3.2. Community detection, for instance, bears an advantage against traditional clustering methods in that it allows for a non-predefined number of clusters and cluster overlap [Papadopoulos11]. As an overview, community detection retrieves the most densely interconnected sets of nodes in a user-tag-resource social network graph [Newman03].

2.2.4 Association rule mining

To this point, we have presented the ways by which semantic user preferences can be retrieved and updated. In an intelligent preference learning environment, it is particularly significant to discover the specific correlations and rules that govern user preferences. These rules illustrate specific *behavioral patterns* of an individual user or a group of users and are as important for meaningful decision making as is the collection of conceptual user preferences, if not more. Such semantic patterns can be mined in the form of association rules. An association rule is generally understood as the relation between conjunctions of attribute-value pairs (categories) called antecedent and consequent.

The term association rule was coined by [Agrawal93] in connection with his proposal of the apriori algorithm in the early 90s. The idea of association rules was later generalized to any data in the tabular, field-value form. There are two basic characteristics of an association rule -- support and confidence. These are called interest measures and a minimum thresholds on these values are the key externally set parameter of an association rule mining task.

The expressivity of rules on the output of the association rule learner is particularly important for multiple reasons. For example, the widely accepted standard a priori algorithm does not allow using disjunctions and negations in the rule, which makes it cumbersome to incorporate negative user feedback due to the lack of negation connective and results in an excessive number of rules, which could be substantially reduced if the disjunction connective was available.

Another concern when selecting the association rule mining algorithm is the list of interest measures supported. The selection of the interest measure(s) fundamentally impacts the quality of the rules on the output of the association learning task as judged by a domain expert. A survey of interest measures is given e.g. in [Zhang09]. Constraints involving multiple arbitrary interest measures are supported by the arules package [Hahsler05].

There are multiple algorithms/implementations supporting some of the advanced features. For example, the R package arules [Hahsler05] allows to use multiple arbitrary interest measures. A disjunctive association rule mining algorithm is described in [Nanavati01] and negative association rule mining algorithm in [Antonie04].

To the best of the authors knowledge, the GUHA ASSOC procedure as implemented in the LISp-Miner system is the only single association rule mining algorithm with a publicly available implementation that offers a full range of logical connectives. This algorithm also allows constraining the generation of basic and derived boolean attributes separately for each cedent thus giving the possibility to significantly limit the search space. A cedent is a group of attributes which can be used anywhere in the antecedent or consequent of the rule.

Considering its broad scope, including more than forty years of research, GUHA ASSOC seems to be a suitable candidate for a rule mining algorithm in the LinkedTV project. Also, as perhaps one of few association rule mining algorithms, LISp-Miner offers a web service interface through the SEWEBAR-CMS project [Kliegr11]. The language used for description of mining models is an XML exchange format based on the standard Predictive Modeling and Markup Language [Kliegr10].

2.3 Profiling in digital media

In personalised digital media systems in particular the metadata desultory and sparcity in heterogeneous multimedia content is the prevailing challenge in understanding and capturing user preferences. In the approach of [Martinez09] for a personalized TV program recommendation, the system relies heavily on explicitly provided user preferences, while avoiding to delve too deep into automatically extracting fundamental user knowledge. In addition, while further preference analysis employs explicit user ratings of viewed content, the system also implicitly takes into account hidden indifference based on unseen content, thus indicating a concrete primitive method to track disinterests.

[Tsunoda08] proposes automatic metadata expansion (AME), a method used to enrich TV program metadata based on the electronic program guide (EPG) data and an associated concept dictionary (ACD). Transactional information such as program recording, program lists browsing, program searching, and voting on viewed programs are processed by a rule-based engine automatically in order to update the user interfaces. Indirect collaborative filtering (ICF) is used to trace peer induced user preferences. Typical users are clustered prior to the similarity calculations of CF.

[Ardissono03][Ardissono04] describe a user-adaptive TV program guide Personal Program Guide (PPG). PPG stores stereotypical information about TV viewer preferences for such categories of users as, for example, housewife. A "general ontology" is employed to model the TV preferences concept space, including hierarchical categories of TV programs from broad concepts like 'Serial' to more specific like 'Soap Opera', 'Sci-Fi Serial', etc. The structured ontological knowledge allows PPG to effectively map different TV program characterizations to single-vocabulary genre categories.

In contrast, [Aguzzoli02] uses free form characterizations of compilations of audio content recorded by the users in a music reproduction platform as features that describe the genre of use of the content rather than its typical musical genre/category (e.g. 'vacation songs' instead of 'rock'). Similarity of these features between recorded compilations in a repository contributed by system users is measured in order to establish a case base used to predict relevant user preferences. The iEPG [Harrison08] "intelligent electronic programming guide" platform utilizes a graph-based organization of content metadata in the multidimensional information space of TV programs, used to structure user preferences and visualize them for the users to access, assess and define nodes of interest, where "me" is the central node of the graph.

In [Zhang05] the authors propose a model computing the level of interest of the user over a Personalized TV Program Recommendation, based on usage history and the actions performed on the player, and also the duration of the action and the consumed content metadata. They compute the Content Affinity, in order to index the level of users' interest to program content, taking notice of users' actions and their associated viewing duration, in order to conclude how the different actions influence of the program to user preferences. The notion is based on the fact that the more one likes a program, the longer time he will spend to watch or rewind it and on the other hand, the less one likes a program, the more time he will fast forward and skip forward while watching a program. An important point is cleared out in [Shin09] where the authors tried to underline the difference between playback duration and the actual viewing time when talking about TV consumption and further classifying user behaviour in accordance with their purpose to consume, store, or approve what they consumed.

The NoTube project[2], which finished in January 2012, aimed to bring Web and TV closer together via shared data models and content. NoTube concentrated on developing services, APIs and applications for personalised television, based on TV content recommendation on the basis of topical closeness with a user interests profile. The user interests profile is generated automatically from the user's social Web activities using the NoTube-developed *Beancounter API*, and the contents of that profile is exposed online and controlled by the user via a dedicated UI for the Beancounter, where the user can select which social Web sources are to be used (e.g. Facebook, Twitter, GetGlue, last.fm), and can add, delete or alter the weightings of selected topics in the profile. The Beancounter monitors changes in the user's social Web activities over time and evolves the user's profile. The user profile in NoTube uses the FOAF vocabulary in the RDF/XML serialization, extended with weighted interests (adding a weight value on the FOAF `interest` property) and published as FOAF-WI[3]. The interests were identified using categories from the DBPedia concept space. Another aspect of NoTube was to extract categories from TV programs using their metadata and

---

[2] http://www.notube.tv

[3] http://xmlns.notu.be/wi/

concept extraction techniques (on their textual descriptions). These categories also use the same DBPedia concept space, and since in DBPedia all of the categories are modeled into a giant taxonomy using the SKOS specification, it was possible to determine similarity rankings between categories based on implementations of concept distance measurement in the SKOS model. These similarity rankings, applied between the set of (weighted) categories attached to a TV program and the set of (weighted) categories attached to a user profile, were the basis for a TV recommendation engine.

2.4 Semantic user profiling

Aiming to leverage the shortcomings of the information stemming from the profiling strategies described in chapter 2.1, to harvest the opportunities of the learning approaches described in chapter 2.2 and in effect to bring together *uniform and meaningful conceptualisations* of the diverse, ambiguous and sparsely described multimedia content pertinent to users, the semantic interpretation of user preferences will be explored within LinkedTV. That said, ontology-based user profiling poses two main challenges: the need for predefined, usually manually constructed, domain-specific knowledge and meaningful semantic mappings between content descriptions and the semantic information in the ontology.

While the term "semantics" is loosely interpreted in personalisation literature as the retrieval of meaningful relationships between content features or user attributes [Zadeh08], this subsection refers to paradigms that employ formal ontological knowledge as the background for building structured semantic user model descriptions in the context of knowledge-based preference learning. A formal ontology consists of: a) *concepts* and their *properties* (which can be subdivided into scalar attributes and non-scalar relations), and *instances* that represent entities belonging to the concepts and b) *axioms* and *predicates* representing the specific rules that depict complex domain knowledge [Weissenberg04].

There are several issues that need to be dealt with in the context of implicitly capturing and representing a semantic user profile: the ontology/ies that can provide a meaningful knowledge base potent to capture domain and user-pertinent semantics (chapter 2.4.1); the expressivity of the ontological language that is most suitable to articulate this knowledge as well as the knowledge learned about the user (user model) (chapter 2.4.1); the means of unobtrusively capturing user behaviour (chapter 2.42); the means to understand user behaviour, i.e. map it to available knowledge (chapter 2.4.3); determining the most suitable representation schema of the user model in a manner that renders the synergy between the model and the background knowledge feasible within the context of an intelligent inferencing engine (chapter 2.4.4).

2.4.1 Ontologies and knowledge representation

Today, vast amounts of information are available on the Web and in other sources. The question arises how this information can/should be used for personalisation and contextualisation. We may assume an ontology forming the vocabulary and world knowledge backbone of the user model which can contain all relevant domain and/or user-specific

concepts and their relationships and that the ontology can provide uniform, compact conceptualizations for ambiguous, synonymous and multilingual knowledge for faster inferencing.

Regarding comprehensive knowledge representation languages that can adequately express an ontology that encompasses information pertaining user preferences, a large number of standards have been proposed by the World Wide Web Consortium. Briskly presented, the RDF[4] data model serves as the basis for assertional languages. It covers basic relational semantics using triples. Within the RDF/S[5] schema basic ontological primitives can be expressed, allowing inference of new triples. Widely used serialisation standards for RDFs are XML[6] and Notation3 (or N3)[7]. Many representation languages support extended semantics based on the RDF schema, e.g. SKOS[8] that expresses more hierarchical information. In regard to formal knowledge representation, languages such as OWL[9] and their logical foundation, Description Logics [Horrocks04] have become prevailing choices [Dalakleidi11] for expressive knowledge representation. OWL's three manifestations (OWL-Lite, OWL-DL, OWL-Full), can be used to achieve the desired trade-off between high expressivity/completeness and complexity. In addition, the extension of OWL, namely OWL 2[10], and its manifestations (OWL 2 EL, OWL 2 QL and OWL 2 RL) can provide additional functionalities[11] (such as richer datatypes, additional property semantics, enhanced annotation capabilities etc). FOAF[12] uses semantic technologies to represent people, their relationships to each other, their interests and activities in a machine-understandable format.

As far as ontologies go, DBPedia [Auer07] stands out as the most prominent organization and concentration of knowledge in the current literature. DBPedia is a shallow ontology that interlinks and semantically structures Wikipedia[13] cross-domain information and constitutes 320 classes described by 1650 properties (as of 8/2011) and ~1.8M instances. It is released in a variety of languages, allowing for multi-language alignment. It is part of Linked Open Data[14] (LOD) initiative that attempts to provide structure to the vast mass of information available online. However the broadness of this information restrain DBPedia to relatively low expressivity, corresponding to the $\mathcal{ALF}$(D) complexity of DLs, namely entailing subsumption between classes and property restrictions [Völker11].

---

[4] http://www.w3.org/RDF/

[5] http://www.w3.org/TR/rdf-schema/

[6] http://www.w3.org/XML/

[7] http://www.w3.org/DesignIssues/Notation3.html

[8] http://www.w3.org/2001/sw/wiki/SKOS

[9] http://www.w3.org/2001/sw/wiki/OWL

[10] http://www.w3.org/TR/owl2-profiles/

[11] http://www.w3.org/TR/2009/REC-owl2-overview-20091027/#Relationship_to_OWL_1

[12] http://www.foaf-project.org/

[13] http://www.wikipedia.org/

[14] http://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData

Partner Eurecom's NERD ontology [Rizzo11] provides a frame for mapping named entities described across several multi-discipline vocabularies and will be a core tool for expanded multimedia annotation in WP2. This ontology can significantly support semantic classification of generic user-consumed content.

Condat's Smart Media Engine (CSME) uses the freely available information in the Linked Data cloud to build and automatically update a semantic knowledge base called "full ontology". It is based on the categories and resources of the German DBPedia. For further information about synonyms and similar phrases the German dictionaries OpenThesaurus[15] and WortschatzUNILeipzig[16] are crawled. The full ontology contains over 2 million semantic objects (URIs) and over 7 million triple relations between these objects. The Analyzer component of the CSME also builds a knowledge subset called "active ontology". It contains all context-relevant information of the "full ontology" according to the current metadata and use case and is employed to reduce the concept space and speed up the content recommendation process.

The Cognitive Characteristics Ontology[17] provides a vocabulary for describing cognitive pattern for users within contexts, their temporal dynamics and their origins, on/ for the Semantic Web. [Belk10] use their OCUM (Ontological Cognitive User Model) as an upper ontology that "*encompasses the core human factors elements for hypertext computer-mediated systems*" and can be reuse to provide enhanced user-centric mappings for personalisation systems.

The NAZOU[18] project provides also user model ontology. Ontology-based user model defines concepts representing user characteristics and identifies relationships between individual characteristics connected to a domain ontology. Such a model is (after its population) used by presentation tools to provide personalised navigation and content. Model can be employed also in content organizing tools (e.g., perform sorting of items based on user's preferences).

The user ontology in NAZOU is composed of two standalone ontologies, which separate domain-dependent and general characteristics: Generic user ontology (OWL) - defines general user characteristics, Job offer user ontology (OWL) - defines characteristics bound to the domain of job offers represented by domain ontology.

There are several ontology languages which can be used as the vocabulary basis for the representation schema of a user profile and many distinct ontologies that can serve as the knowledge base describing the micro-world pertinent to the user in a semantic personalisation system. In essence, the ontology should provide the structure and

---

[15] http://www.openthesaurus.de/

[16] http://wortschatz.uni-leipzig.de/

[17] http://smiy.sourceforge.net/cco/spec/cognitivecharacteristics.html

[18] http://nazou.fiit.stuba.sk/

---

relationships needed to formulate and manage the user model and the knowledge needed to retrieve additional relevant information from this model.

The general purpose, cross-discipline ontologies presented in this section can be considered to provide a core framework for understanding and representing the broad scope of potential user preferences. However, their limited expressivity *de facto* cannot support depiction of advanced information like domain-specific ontologies can. Briskly described, the need to better define knowledge has resulted to a large amount, freely available domain-specific ontologies[19] across numerous domains, available for re-use and alignment within multidisciplinary environments.

2.4.2 Implicit user information tracking

This section briefly describes main measures and patterns of user information, which can be collect implicitly from interaction of users with web and media content and provides an overview of approaches how to track this user information.

Implicit feedback

Implicit user information, also called implicit feedback, is important information to understand user interests. Implicit feedback has the advantage that it can be collected easily without burden of the users, but it is more difficult to interpret and potentially ambiguous. There are many measures, which can express this information based on behaviour of users. Three main sources of implicit feedback exist [Zhang10]: Attention time, Click through, Mouse movements. Other view on measures and behaviour patterns related to the implicit feedback is proposed [Oard98]. There is overlap with the previous set of measures. Observable behaviour for implicit feedback based on three discrete behavioural categories can be seen in Table 2: Observable behaviour for implicit feedback [Oard98].

Table 2: Observable behaviour for implicit feedback [Oard98]

Category	Observable Behaviour
Examination	Selection
	Duration
	Edit wear
	Repetition
	Purchase (object or subscription)

---

[19] E.g. http://protegewiki.stanford.edu/wiki/Protege_Ontology_Library#OWL_ontologies

Category	Observable Behaviour
Retention	Save a reference or save an object (with or without annotation) (with or without organization) Print Delete
Reference	Object → Object (forward, reply, post follow up) Portion → Object (hypertext link, citation) Object → Portion (cut & paste, quotation)

These categories cover all main behaviour for implicit feedback and can help to understand user interest and needs. To successfully using this implicit feedback, it is important to get source information about user – we have to track behaviour of users. Tracking approaches are described in following section.

Implicit information tracking

The information about users can be collected by various techniques. In general, there are many channels how to collect implicit information. This section will provide a summary of approaches related to browsing and activities on the web or media consumption.

Implicit collection does not require intervention by the user. All information is gathered on the background using agents that monitor user activity [Kelly03]. Data can be collected on the client machine (client-side), by the application server itself (server-side), or both. The quality of the collected data available is an overlooked, but critical factor for determining the success of any further processing. Client side monitoring is required to get a precise record about a user interaction with the website [Xu11].

Data acquisition for implicit information analysis is often taken as a synonym for processing server logs. This section shows a range of approaches to collecting usage data and compares them. Some of these approaches are commonly used in practice; some of them are only experimental.

The basic task in the data acquisition (or in some sources data collection) phase is to record visitor actions such as content views, events etc. Additional information may be gathered in order to identify user sessions and assign weights to the actions.

Table 3: Implicit User Information Collection Techniques [Gauch07]

Collection Technique	Information Collected	Information Breadth	Pros and Cons
Browser Cache	Browsing history	Any Web site	**pro:** User need not install anything. **con:** User must upload cache periodically.
Proxy Servers	Browsing activity	Any Web site	**pro:** User can use regular browser. **con:** User must use proxy server.
Browser Agents	Browsing activity	Any personalised application	**pro:** Agent can collect all Web activity. **con:** Install software and use new application
Desktop Agents	All user activity	Any personalised application	**pro:** All user files and activity available. **con:** Requires user to install software.
Web Logs	Browsing activity	Logged Web site	**pro:** Information about multiple users collected. **con:** May be very little information since only from one site.
Search Logs	Search	Search engine site	**pro:** Collection and use of information all at same site. **con:** Cookies must be turned on and/or login to site. **con:** May be very little information

Many studies are based on browser cache, proxy [Aquin10] [Fujimoto11], desktop or browser agents [Xu11]. Hybrid approaches are available too. All approaches were compared, but there is no clear answer, which is more or less accurate [Gauch07]. On

Table 3 the main pros and cons of all approaches [Gauch07] are mentioned. In general, the approach that does not require any additional software on the client side is preferred.

One of the compromise solutions is using a proxy server [Aquin10], [Fujimoto11], [Fujimoto11b]. There is need only one setup at the beginning and all interactions of user are collected at the proxy server [Holub10]. The drawback can be in identification of users. Group of users can use the same proxy server and there is a problem to distinguish them. This solution is very often limited on a single computer. However, there are solutions to use login information and install proxy on various numbers of computers.

The approach based on browser cache [Jakobsson08] has no limitation on web sites, but it needs to send out cache information periodically.

Third category is based on agents. Desktop agents are implemented as standalone application and browser agents are implemented as plug-in to an existing browser [Ohmura11]. The main disadvantage is that user has to install additional software. The advantage is in possibilities and intelligence of this solution. Agent has much more available information about user interactions and can provide better browsing assistance (fill out forms, highlight content, modify content etc).

The logs related category (Web and Search Logs) does not require additional software. Although it is prone to collect much less information than agents, it is one of the most used approaches. There are two main sources: browsing activity and search interactions. Details about browsing activity logging will be described further here. Search interactions can provide information about queries and help to collect information about user [Ghorab09], [Park11]. The drawback is in collecting information related only to search interactions.

Methods for User Identification

User identification is a crucial ability for all of previously mentioned approaches. There are five basic approaches to user identification: software agents, logins, enhanced proxy servers, cookies, and session ids [Gauch07]. Cookies are the least invasive technique and are widely used.

Software agents, logins and enhanced proxy servers are more accurate. The drawback of these approaches is need of user participation. The general characteristic is to register user and log in to the system.

Cookies and session ids are less invasive. This method is unobtrusive and there is no need of user participation. The drawback is in different sessions across different client for the same user or one shared session for multiple users using the same client.

Client-Side Tracking

The most well known form of client-side tracking on the web uses javascript tracking code, which is a part of the website. Tracking is executed on the client side and has much more information about user interactions, but usually the process is absolutely unobtrusive for the user, who does not know that he is tracked and, particularly, does not need to install any tracking software.

Although client-side tracking is deemed more convenient and can provide far greater granularity than log-files, the server side tracking has not yet completely lost its place.

Server Side Monitoring

Server log-files are the most commonly known type of server side monitoring. This was a primary approach to analysis in the past. It can still provide a useful insight, although it is becoming obsolete. Server-side monitoring is not, however, a synonym for server log-files. A server-side application may log any interaction (or its implications) with the visitor, which is propagated to the server. Some approaches use an application-level server side tracking meaning that the server application is responsible for the generation of content is also in charge of tracking.

For example, when the server application generates the content, it can record which pieces of information it contains. An authoritative survey [Facca05] marks application-level server side tracking as *"probably the best approach for tracking Web usage".*

Custom server side monitoring has several important virtues, when we consider its deployment in a dynamically built website:

1. Unobtrusiveness: no demands on the client, including scripting support.

2. Invisibility: the client can't know (in principle) whether he is tracked and what data are collected.

3. Speed: the data are readily available on the server-side, no additional client-server communication needed.

4. Interlinked with content: all information in the underlying database, from which the content is generated, is available for tracking purposes.

Despite these advantages, custom server side tracking has apparently never gained much popularity. It suffers from the same accuracy problems (what concerns the omission of hits) as log based solutions. Foremost, handling personal data on the server side seriously compromises user privacy.

Alternative approaches to Client-Side Tracking

The top-of-the-line approaches need a special hardware to perform eye tracking in order to locate the exact part of the page the visitor looks at. A relevant research to e-commerce Web Usage Mining (WUM) using eye tracking is presented in [Granka04]. Other approaches utilize a standalone software or a browser plug-in to track user interaction with a website. According to [Barla06] this is the case with research presented in [Thomas03]. Client-side tracking software is often used for usage pattern analysis, because it gives a possibility to gather very detailed information, which would not otherwise be accessible. This information is used mainly for the following types of tasks (according to [Kim06]):

- Recognition of common areas in a Web page using visual information - based on a visual analysis of the page as seen by the visitor in a browser.

- Vision-based Page Segmentation - idea behind vision based page segmentation is to mimic the way how humans visually process web pages.

2.4.3 Semantic classification

Semantic classification addresses the problem of understanding the substance of a content item (e.g. multimedia metadata, tags, keywords, articles, short textual descriptions) pertinent to the user. The main challenge underlies in interpreting data for which no, little, or vastly heterogeneous descriptions are available and expressing it in a uniform, machine-understandable format, i.e. mapping the content and its impact to the user to available ontological knowledge. Typically, this task requires text processing to extract the textual descriptions of the content and to recognize entity types, as well as a mapping mechanism to detect correspondence between the terms and semantic entities. Although media content is expected to be adequately interpreted within the LinkedTV context (WP2), free-form extra content relevant to a user can not be expected to be semantically pre-described, especially not in a common vocabulary. Evidently, addressing this challenge lies in the borderline between WPs 2 and 4 and is subject to further alignment of cross-work package responsibilities. Nonetheless, we find it important to consider several multidimensional methodologies suitable for LinkedTV to align diverse multimedia descriptions with textual descriptions, which can work in synergy with WP2's Linked Media Layer.

[Cantador08] employs advanced lexical filtering to create a common vocabulary of social tags for enriching user profiles by grounding tags based on Wikipedia[20] and the Wordnet lexicon [Miller, 1995]. This vocabulary is subsequently used to map the tags to concepts based on reference ontologies. Their filtering requires intensive term pre-processing to unify tags with concepts, yet abstains from including other relational information relevant to a concept. Similarly, [Gemmis08] induce extensive linguistic prepossessing to semantically index documents. The classification process includes disambiguation of polysemous words based on their context and relies heavily on natural language processing techniques such as Word Sense Disambiguation and on the WordNet lexical ontology.

Condat's Smart Media Engine [CSME11] tool analyzes the semantic content of metadata automatically and enriches metadata with freely available information from the Internet. Hence, content can be categorized and compared without any medium borders. Out-of-the-box structures like TV or radio programmes and websites are supported. The imported text data is analyzed by the TreeTagger[21] POS tagger tool, with various corrections and enhancements made by Condat, and mapped into the ontology. Every metadata object gets a list of weighted semantic objects (persons, places and other significant phrases), the so-called "Semantic Fingerprint".

---

[20] www.wikipedia.org

[21] http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/

In [Trajkova04] and [Sieg07] concept vectors are constructed for each ontology concept by indexing a training set of web pages in order to classify an ontological user profile. In these approaches, the user profile is a bag-of-words that are mapped to the ontology, implying a taxonomy-based profile. There is no accounting for relations between terms.

This paradigm was extended by [Tsatsou09] [Tsatsou11] where domain analysis of unstructured textual data via lexical graphs produced loose semantic correlations between domain terms. Terms correlating with available ontology concepts were used to populate vectors, called *concept vectors*, which enriched the concepts, accompanied by the confidence degree by which the term relates to the concept based on graph statistics. The concept's synonyms (if concept type: noun) or variations (if concept type: named entity) were also introduced in the concept's vector with a strong confidence degree. This method provided more substantial and compact semantic information about the concepts than the aforementioned work, producing an enriched ontology in which each concept was complemented with descriptive information that could be used to recognize this concept in a given textual description of an item, even if the explicit concept string representation was not present in the text.

Consequently, the enriched ontology was used to detect fuzzy semantic mappings between raw textual content descriptions and the ontology concepts, thus semantically interpreting user transactions and producing high-precision concept classification. Profile concepts could be restricted by ontology properties via heuristic examination of their domains and ranges. The advantage of this method constitutes of the high quality of the classifications, especially concerning named entities. The disadvantage consists of the need to pre-train lexical graphs to analyse the relations between terms in a given domain, although the analysis was far from exhaustive and the graphs could effortlessly be dynamically adapted to domain changes and collect mass intelligence from aggregated user generated content.

While this method is yet to be extended to support more intelligent relation classification, [Zhang07] provides a more sophisticated method to extract relations between concepts pertinent to the user through the creation of a semantic log where usage patterns are mined by examining the ordered paths between classified interests in the user's transaction history.

2.4.4 Semantic user modelling

The main goal in identifying the most suitable user modelling schema is to determine the syntax that adequately represents the user preferences and their relations while supporting context distinction and expressing the level of preference.

As aforementioned, a considerable part of the literature uses an instance of a reference ontology or the ontology in its entirety to activate interesting concepts while determining how impactful they are for the user [Sieg07]. Similarly, yet more compact techniques construct personal user ontologies as subsets of the reference ontology, as in the case of [Zhang07] where a directed *concept-graph* is modeled for the user, with its nodes denoting ontological concepts and carrying a vector of instances that belong to the concept along with the confidence degree of that inclusion. The graph's edges depict semantic relationships

between the concepts while incorporating statistical and temporal information about the relevance of these relationships to the user.

In [Zeng10] a formal definition of E-foaf :interest Vocabulary Specification[22] was presented for describing user interests. The vocabulary is based on RDF/OWL, Linked data idea and is aimed on extending the FOAF vocabulary. The interest can be interoperable across various applications. The vocabulary was evaluated on the search application on the Web. This representation is able to describe time evolving interests and their relations.

However, these approaches pose three main hinders (without all three pointers necessarily applying to the entirety the aforementioned approaches): i) scalability problems, especially with respect to the diverse and broad preferences emerging in a multimodal linked media environment, ii) inability to express complex patterns in user behaviour, most prominently in different contextual situations where different relations might pertain the same concepts and iii) inability to express disinterests.

Other approaches use a preference set where extracted user interests and their importance degree are collected in a unified vector space as in [Li02]. [Kearney05] constructs an ontological user profile in the weighed sum of the total of pre-classified user visit profiles history. Again these approaches lack the expressivity to articulate behavioural patterns.

Consequently, other paradigms construct rule-based user profiles to address these limitations. In [Tsatsou09][Tsatsou11] the profile is formulated in a disjunctive DL axiom, containing quantified weighted preferences and primitive relations connecting them, as well as conjunctions of frequently correlated preferences. The advantages of this approach constituted in the lightweight semantic descriptions which were stored and efficiently, potent to be used for inferencing directly on the client, thus promoting privacy-preserving decision making, while the use of uncertainty weights and confidence degrees in the semantic descriptions and the user preferences allowed for fuzzy reasoning in order to produce ranked recommendations.

2.5 Summary and conclusions

In the overview of previous work, we have examined the possible sources of information used for user profiling and the main profiling approaches that take advantage of these sources, we have narrowed our search to salient preference learning strategies that exploit the capabilities of the profiling strategies in order to elicit complex knowledge about user behaviour, we have identified key aspects of profiling approaches in digital media environments and consequently focused on the analysis of specific techniques for semantically capturing and modeling user preferences.

The reasoning for selecting semantic user profiling as the predominant strategy for personalizing user preferences in LinkedTV origins from the need to take advantage of all the

---

[22] http://wiki.larkc.eu/e-foaf:interest

various sources of information deriving from heterogeneous resources and extended user activities (player click-through behaviour rather than mere content consumption) in a networked media platform while leveraging their opportunities and challenges. To this end, semantic user profiles can sufficiently address the intricacies posed within networked media environments by reducing the dimensionality and scale of heterogeneous media data with an interest in keeping user models lightweight and efficiently meaningful, while exploiting the capabilities of the described machine learning mechanisms to their fullest. To defend this reasoning, we must briskly examine in retrospect the main issues addressed in this chapter.

In a personalised environment information about the user can be either explicitly defined by the user himself or implicitly extracted through his interaction with the system and relevant resources. We must assume that within the LinkedTV platform a user will be able to directly define or manipulate his profile. However, due to the obtrusive and outdated nature of explicit preference acquisition, we cannot rely on the user to provide this information. In effect, the focus of the LinkedTV personalisation task will lie on implicitly extracting user preferences while addressing the privacy compromising hampers that implicit preference acquisition poses. The decision of the appropriate implicit information tracking strategy for LinkedTV involves balancing off the attributes of known tracking techniques in order to retrieve information in the least invasive manner that can provide the most meaningful information.

The sources of implicit information pertaining a user are three-faceted, involving data extracted from the content a user consumes, similarities with peers in a social information exchange environment and predefined knowledge encompassed in ontological structures. These aspects come with advantages and hampers, pinpointed in the previous chapter, the main being the cold-start and scalability problems for the content-based and peer-based approaches while knowledge-based approaches are hindered by the lack of mappings between free-form information and the ontological knowledge. A hybrid approach taking into account both content and peer-based information while understanding and aligning them under uniform ontological conceptualisations, thus capitalizing on available knowledge, is esteemed as the optimal trade-off.

In addition, expected information overload, instigated by the diversity and vastness of information in multimedia environments, can be efficiently managed by the representation of the information in compact, lightweight ontological conceptualizations. These issues have motivated us to explore an approach to extract the semantics of user behaviour with respect to the viewed and consumed content and his interactions with peers and express them in a uniform ontological vocabulary. This uniform vocabulary will further aid in obtaining more meaningful knowledge about the user based on observation of his transactional behaviour, namely learning relations about user preferences and learning relations about the domain from aggregated user information.

In conclusion, specific aspects of semantic user profiling within the scope of LinkedTV such as appropriate reference vocabularies (ontologies) and knowledge representation schemata have been examined, along with potential means to extract semantic information via classification of raw data to ontological knowledge. Semantic classification is deemed

necessary in WP4 in order to uniformly understand user preferences pertaining the transaction of the user not only to the platform's expressively annotated media content (following the work of WP1 and WP2) but also to the heterogeneous linked content that the user might consume in complement to the media. Finally, an overview of potential representation schemata for a semantic user profile has been conducted, concluding to the most advantageous schema for LinkedTV, which summarises the general theme of the personalisation task: *say less, mean more*.

3   State of the art in contextualisation

Context is a difficult notion to capture in a recommender system, and the elements that can be considered under that notion are manifold: user tasks/goals, recently browsed/rated items, social environment, physical environment, physical reaction to presented content, and location, time, external events, etc. This section will provide an overview of the state-of-the-art in contextualisation and conclude on the best practices to be extended within the LinkedTV platform. Two contextual influence factors are going to be explored: on-platform interaction with the content, taking also into account concrete user situations (e.g. time, location) (chapter 3.1) and behavioral tracking in reaction to the content (chapter 3.1). Finally, an overview of known methodologies to capture and express user context in digital media environments will be presented (chapter 3.3), with an interest in identifying the level of involvement of context in existing personalised media approaches and the particular challenges that need to be addressed within LinkedTV.

3.1   Transactional context learning

The majority of existing contextualisation approaches rely on segmenting user behavior or filtering content based on factors like the time, location and company of other people [Adomavicius08]. Beyond this semi-explicit information, the most straightforward contextual influence obviously comes from recent content consumptions: they define the most recent context and are typically a strong guide for predicting content recommendation, especially pertinent to media consumption [Bazire2005].

[Sieg10] illustrates the importance of distinguishing long-term preferences and short-term context related preferences in a dynamic personalisation environment, with the short-term preferences stemming from recent content consumption. The approach achieves a trade-off between accuracy-achieving factors, like continuously making 'safe' recommendations of the most prominent long-term user preferences and diversity, the lack of which might significantly reduce recommendation performance.

Similarly, the problem of tedious repetitive recommendations was addressed in [Abrams07], where a novel feature was taken into account on contextual personalisation of ads: the *ad fatigue* factor, which refers to the repetition of recommended ads in frequently visited web pages, such as a user's home page. The ad fatigue factor is used to diminish the impact of such tiresome recommendations and refresh the user profile.

However taking into account only the current or recent transactions in a user session poses two limitations: a) the contextual frame is too narrow to take advantage of the full knowledge the system has for the user and the domain and b) the systems might lack sufficient data about the behaviour of the user due to the limited features of the current content. To this end [Adomavicius08] generalization of the context is opted.

[Adomavicius08] identify three methods to apply contextual information in personalised environments. The first two address direct context-based filtering while the third takes into

account the need to recognize contextual user preferences: a) contextual pre-filtering, where context information is used to select content data prior to personalisation, b) contextual post-filtering, where personalisation takes place first and results are further re-filtered and c) contextual modelling, where the user profile itself is contextualised.

[Palmisano08] elicit hidden contextual information from data by identifying latent variables for specific sessions, used to identify patterns in the user behaviour regarding specific topics, able to recognize contexts without having to map these contexts to a specific user in a multi-user environment. In addition, [Weissenberg04] uses ontologies to provide location-aware and situation-aware personalised services. The authors identify a main contextualisation issue that goes beyond recent/current user status: the need to identify specific user contextual situations within which the features that characterize them remain invariant in certain time periods.

It can be concluded that while the need to specify user preferences with respect to the content at hand and make user profiles more agile to circumstantial changes, significant benefits in intelligent user behaviour understanding lies in recognizing persistent preference patterns for certain contexts.

3.2 Reactive behavioural tracking

Behavioural tracking technologies generally regroup all means to observe, analyse and process non-verbal behaviour of a user. This is a very large topic spanning from speech (audio) analysis and gestures understanding to emotions reading, mouse clicks or navigation history. The possibility of automatically understanding human behaviour has already vastly been explored. In [Vinciarelli09], one may find a very good survey on social signal processing and behaviour tracking for non-verbal communication. In this section, we will focus on audio-visual features and the behavioural tracking which falls under the scope of computer vision technologies. Moreover, we focus on TV or home related experiences applications where the state of the art is much less deployed.

Behavioural tracking will focus on implicit behaviour analysis of the user physical reaction to the presented content, which means that explicit gestures for interfaces control are not the scope of this section even if the technology used for behavioural tracking can of course provide information about e.g. explicit gestures recognition. User behaviour can be extracted in various ways from different kinds of input data.

With the availability of cheap cameras which are able to acquire both the classical RGB data, but also the depth map representing the distance of each pixel from the camera, vision-based behavioural tracking has made a huge progress. Before those cameras were available behavioural tracking was made by a few research groups using expensive cameras like time of flight cameras or 3D capture systems implying several cameras and infra-red markers, like [Vicon]. Other research groups worked on more classical sensors like simple cameras or stereo cameras, but those devices needed a lot of complex algorithms to provide less precise information and ineffective results for real-life environment and applications.

The Microsoft Kinect sensor is a cheap and effective device which also gained popularity within a large public with its explicit gesture analysis for video games such games for Xbox[23]. Also other 3D camera-based explicit interfaces developed for TVs and interactive adds [SOFT][INTEL12] are more and more popular and lead the public to be more and more aware about these technologies for home and TV-based applications. This trend already pushed some TV manufacturers like Samsung to propose new cameras directly embedded into the TVs [SAM12]. Even if those efforts mainly intend to provide explicit control on TV interfaces, the same systems can also be used in implicit interfaces and behavioural tracking. Moreover some implicit data like the sex or age are already inferred from the cameras.

Within the work on implicit interfaces we can find the work of Microsoft research on the Kinect to extract face emotions and provide them to avatars as it can be seen in Figure 2 where the face emotions are extracted in real time (right image) and then simulated on the left image avatar [MIC_AVA12]. Figure 3 shows several avatars together mimicking real users face emotions. Interfaces showing data (here music) use real-time context as social networks or related songs to provide more context-aware information (Figure 4) [MIC_AMB11].

Figure 2: Right image : face and emotion tracking, Left image : emotion reproduction on an avatar. Extracted from [MIC_AVA12].

---

[23] Xbox video games: http://www.xbox.com

Figure 3: Several avatars who reproduce in a virtual world what their corresponding users are performing in real world bu in separate locations. From [MIC_AVA12]

Figure 4: Context-based information display for a music playlist. From [MIC_AMB11]

Some interesting implicit interfaces were developed using the notion of proxemics in multi-user environments. The proxemics theory was introduced by anthropologist Edward Hall in [Hall66]. The idea is that people's physical distance is correlated to social distance. He noticed several areas surrounding people that suggest certain types of interaction:

- Intimate distance (0 to 45 cm): a really close space with high probability of physical contact, you can feel heat and odor from another person. It's a distance for touching, whispering or embracing someone. It indicates a close relationship like with lovers or children.

- Personal distance (45cm to 1.2m): distance for interacting with relatives like family members or good friends. The other person is at arm's length, only ritualized touch can happen like handshake. Unrequested penetration of this space will provoke discomfort, defensive postures and even avoidance behaviours.

- Social distance (1.2m to 3.5m): distance for more formal or impersonal interactions. It's the distance you naturally pose when you meet stranger and establish a communication process with them.

- Public distance (3.5 to infinity): distance for mass meeting, lecture hall or interactions with important and well-known people.

In the field of proxemic interaction proxemic relationships between people, objects, and digital devices are used together. The design intent is to leverage people's natural understanding of their proxemic relationships to manage the entities that surround them.

[Greenberg11] identified five essential dimensions as a first-order approximation of key proxemics measures that should be considered:

- Orientation: the relative angles between entities; such as if two people are facing towards one another (interpersonal orientations).

- Distance: the distance between people, objects, and digital devices; such as the distance between a person and an interactive display (interpersonal distances).

- Motion: changes of distance and orientation over time (orientation and distance variations in time).

- Identity: knowledge about the identity of a person, or a particular device.

- Location: the setup of environmental features; such as the fixed-feature location of walls and doors, and the semi-fixed features including movable furniture (this includes 3D reconstruction of the scene).

Hello Wall [Streitz03] and Vogel's public ambient display [Vogel04] introduced the notion of "distance-dependent semantics", where the distance of a person to the display defined the possible interactions and the information shown on the display. The space around the display is separated into four discrete regions having different interaction properties. [Ju08] developed a proxemic-aware office whiteboard which is able to switch between explicit (drawing on the whiteboard) and implicit (data display) interaction depending on user's position.

[Ballendat10] developed a system which activates when the first person enters, shows more content when approaching and looking at the screen, switches to full screen view when a person sits down, and pauses the video when the person is distracted. Figure 5 which is extracted from [Ballendat10] shows the different displays on the TV during the observer arrival. This system was initially designed to be used with precise motion capture systems [Vicon], but part of it works also using low-cost depth cameras like the Kinect sensor.

As it can be seen here, the systems applying implicit interaction and behavior tracking in TV setups are existent by limited in number. As the market (both for depth cameras and TV manufacturers) directs towards explicit and implicit interaction based on people behavior we propose here to develop a system which is fully adapted to TV setups and which integrates with other profiling technologies.
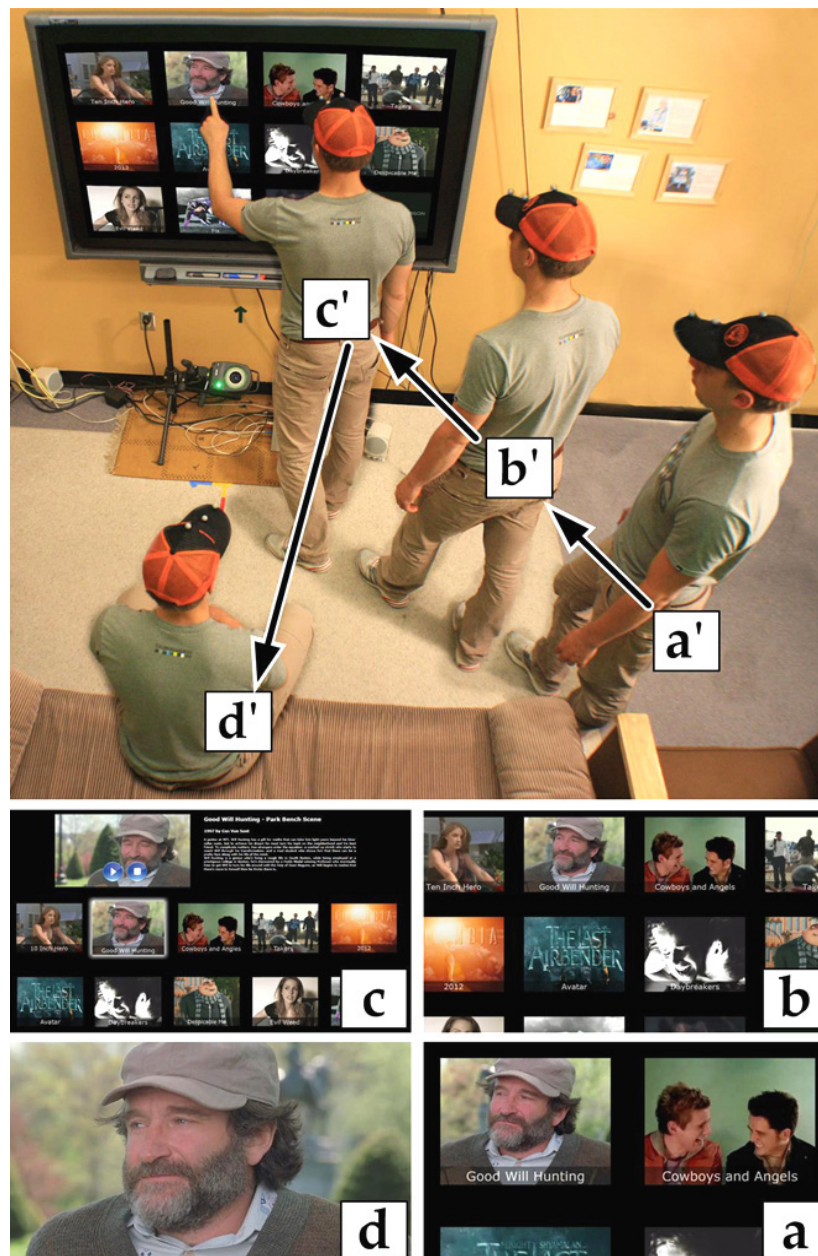
Figure 5: Proxemic Interaction: a) activating the system when a person enters the room, b) continuously revealing of more content with decreasing distance of the person to the display, c) allowing explicit interaction through direct touch when person is in close distance, and d) implicitly switching to full screen view when person is taking a seat. From [Ballendat10].

3.3 Contextualisation in digital media

In generic personalised multimedia applications, the main challenge in eliciting contextual information lies in understanding and aligning multimedia content, underlining the need for multimodal context mining. The problems pertaining the convergence of TV and the Web are drawn by the dynamic nature of the web, which has opened a whole new perspective by offering the possibility to interlink behaviour with respect to multidisciplinary content [Martinez09].

In many media-oriented platforms, the recent viewing context itself is the sole factor considered to provide targeted recommendations to a user, where audiovisual content similar to current display is presented to the user, regardless of his specific preferences. For instance, in GMF4iTV [Huet05], users are able to interact with TV programs through active video objects. The user may select visible objects using a regular remote control, or a PDA and associated additional content can be displayed. In other applications like InfoSip [Dimitrova03] this process is called Content Augmentation and is based on current play context: InfoSip can support answering the viewers' most frequently asked questions using a remote control, such as who, when, where, and what, by interacting with the visual content.

Older studies on TV navigation systems have already indicated that the types of programs desired by a viewer vary according to the specific time of the day [Isobe03]. In [Martínez09], implementing a personalized TV program recommendation system the system indentifies the periodical habits of the user, so that the recommendations target the users' leisure time or time available for media consumption. Profile information is enhanced with supplementary information such as demographics and lifestyle such as age, gender, profession, their TV schedule or leisure time.

[Taylor02] has indicated that the users' attention while watching TV and even their interest is segmented into three levels of viewers' engagement according to time and availability. The first level pertains to the user arriving from some sort of activity or work and therefore pays the minimum attention, the second level comprises the medium engagement, concerning programs of general interest like the news, and the third level concerns the best level of engagement which is estimated by prior training. The authors correlate every part of the day with a level of engagement; for example, during the mid-evening period, the levels of engagement vary, while higher levels of engagement are between 8.30 and 9pm.

In NoTube[24] the authors take into consideration the following context factors: current location, time of the day (dinner time, evening), day of the week, time of the year (summer, winter), activity (traveling), device and multimodal capabilities, social settings and also moods and feelings. They use them to extract the user preferences that are relevant to the context from a user model. They use layers to describe different knowledge domains of a user model, i.e. temporal, spatial, geographic, music-specific, movie-specific. These are separated in hardly changing, slowly changing and quickly changing (context) parts.

[Song12] highlights that context related to IPTV services, should be classified into user context information, such as the identity, location, preference, activity and time, the device context information, such as the screen, the supported content format and the terminals' location, the network context information, meaning the bandwidth and the traffic condition, and finally the service/content context information, which in interactive television is divided into the content description, the video objects and the program interaction.

---

[24] www.notube.tv

The combination of user and technical context is also considered in mobile TV environments. Mobility allows the terminal to change location without service disruption, incorporating situational context information. CoMeR [Yu06], a context-aware media recommendation platform, underlines the importance of utilizing both facets of user-pertinent information: the long-term user preferences and the context. Given the challenges posed in resource-constraint platforms they consider two distinct of context: the user situation context and the device capability context represented based on an ontological structure in OWL. The user's situation is defined by his location, activity, and time, and the media terminal's capability is defined by the device's operating capacity, its display characteristics, the network communication profile, and the supported media modality.

3.4 Summary and conclusions

As concluded from the aforementioned overview in existing work the notion and the facets of contextualizing information pertinent to the user is still quite vague and ambiguous. In linked media in particular, context is still heavily focused on analyzing the current or recent content, expanding its features to more comprehensive descriptions. While this augmentation of the audiovisual content is particularly relevant to LinkedTV, where triggerable video objects are connected to additional content and information, however can only consist of the starting point for extracting richer contextual information and adapting them over different situations of a user rather than just to a specific aspect (subdomain) of the concept space.

The most prominent factor in context-aware environments is the spatio-temporal dimension of the user. Time (also distinguished in several facets such as concrete time of the airplay, seasonal time, purpose time – leisure, work etc) and location have served as panacea to systems aiming to understand more about the situation of the user but little work has been conducted towards capturing additional circumstances that might affect user behaviour.

With the coalition of behavioural tracking strategies, LinkedTV aims to go a step further from existing work by processing broader contextual information, which will go beyond the usual factors such as content, location and time and take into consideration contextual attributes such as the activity, status, attention and mood of the user based on his physical behaviour with respect to viewed and recommended content. Several approaches such as detection of the user's facial expressions, the distance from the client device and distance from actors in his social environment will serve towards implicitly extracting the user's concrete situation.

Obviously the main challenge of the contextualisation task will be to align this diverse contextual information and determine how they can be used in conjunction with user preferences. The synergy between context and long-term preferences still remains a poorly-addressed issue in targeted content recommendation. Intuitively, in an information system like LinkedTV aiming at a "say less, mean more" personalisation strategy which requires lightweight information modeling and fast and efficient inferencing, it is conceivable that out of the three possible uses of context information (pre-filtering, post-filtering, contextual modeling [Adomavicius08]) the latter is the optimal contextualisation paradigm.

4 User profiling requirements & specification

Following our choice to take advantage of the expressiveness and structure potentials of semantic user profiling, that can be used to uniformly encompass meaningful, low-dimensional information as derived from chapter 2, this section is going to provide a requirements analysis and first specification on the profiling decisions and workflow within the scope of LinkedTV. These decisions comprise taking into account: how the user preferences are going to be unobtrusively captured (retrieved and semantically interpreted) and learned, how explicit information will be intertwined, the schema with which preferences are going to be represented, the semantic knowledge upon which the classification and schema are going to be based, the specific user needs and LinkedTV scenarios' demands that the proposed approach plans to address and the storage and communication strategies that the profiling task will follow with respect to the platform requirements and towards safeguarding user privacy.
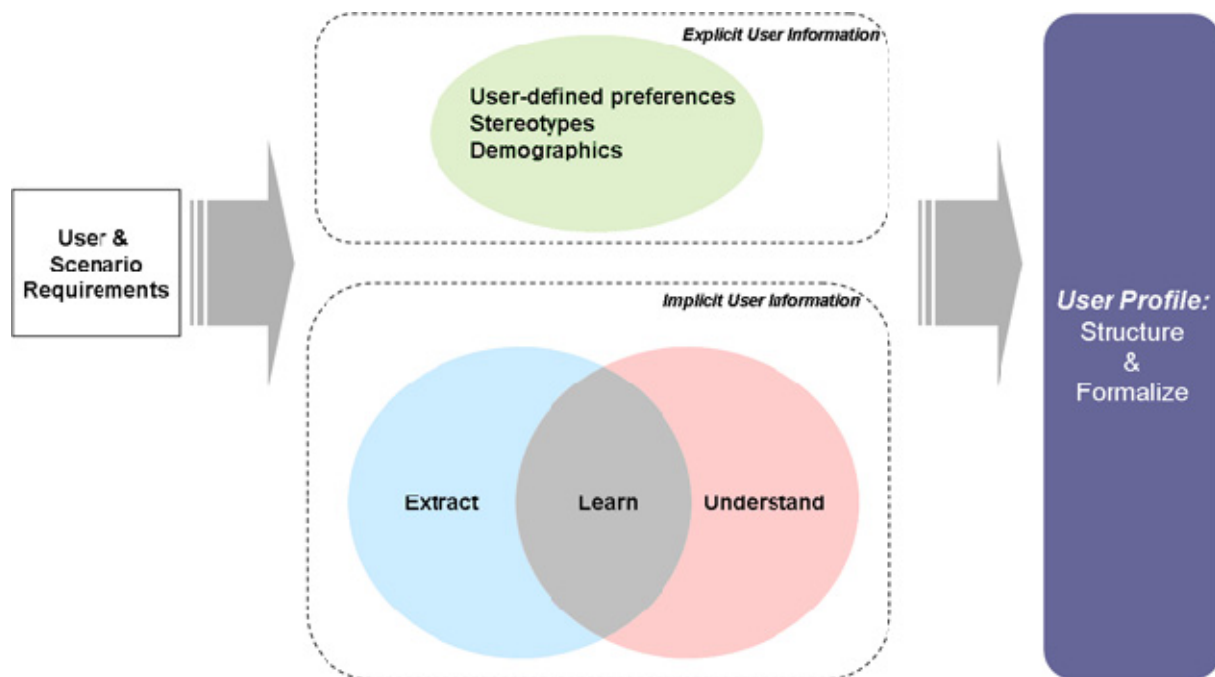
Figure 6: The LinkedTV user profiling methodology

4.1 Functional requirements

Based on the different components of the LinkedTV platform, there are several distinct needs that the profiling task needs to address, especially with respect to the project's use cases (WP6) and the modalities that the user interface (WP3) and the platform's media player intends to incorporate. Three use case scenarios are considered within LinkedTV: a news scenario (RBB), a cultural heritage scenario (S&V) and a media arts scenario (UMONS) with distinct requirements about what they expect from the personalisation system to capture about the user and what recommendations are expected to be made. Furthermore, the personalisation process is expected to be able to understand and align information derived

from the media content (WP1, WP2) and from heterogeneous external content that the user is expected to complementary consume.

In addition, there are several requirements/information that the profiling task expects the platform to be potent to address besides the information about the media content, such as recognition of the user's actions on the interface and what information about the extra web content can be obtained.

Requirements to address

- Unobtrusive and seamless preference extraction and learning with no user involvement required.

- Lightweight modeling of user preferences and background knowledge, aiming for fast and efficient personalisation.

- Ability to factor demographic information and explicit preferences - ability to refactor the profile based on explicit feedback (e.g. rejection of a recommended item, item re-ranking)

- Multilingual support: the target audience and available content from our three scenario partners pertains to three different languages: German, Dutch and French.

- Interoperability between multimedia content knowledge and general domain knowledge.

- Ability to handle the diversity in the heterogeneous information pertinent to the user based on both interaction with the platform's provided content as well as with external resources (social networks, Wikipedia, miscellaneous web sites).

- Interweaved information management of user behaviour, spanning between transactions with the media content (annotations), linked media descriptions and external consumed content.

- Ability to assess user transaction activities on the media player (bookmarking, skipping, pausing, replaying etc).

- Preserve user privacy.

Prerequisites

- Designation of what information can be made available from external resources (e.g. social network profile, peers interaction, web page text etc)

- Designation of the types of transaction activities offered within the player (bookmark, skip, save etc).

- Definition of the information available from the Linked Media Layer.

- Storage and processing capacity within the client

- Storage, processing and communication capacities within the server.

4.2 Semantic knowledge base

In a linked media environment, the user is expected to interact with all kinds of content, spanning from the multimedia content itself (with information varying from visual features, the audio transcripts and media annotation) as well as a wide variety of miscellaneous content relevant to the media (news articles, program summaries, bios, ads, encyclopedic information, photos etc). It is significant for the efficient elicitation of user preferences to have a uniform and compact vocabulary to classify this information under. To this end, ontologies provide the needed expressivity and conceptual basis.

The requirements on deciding over the most suitable semantic knowledge for the users of LinkedTV includes determining the level of granularity, the semantic precision, and the expressivity of the user-centric ontology with regard to appropriate inferential services.

A core ontology aiming to adequately describe domain knowledge relevant for a user in a heterogeneous hypermedia environment is expected to be rather broad. It needs to cover anything from the description of the visual features, to the high level topic conceptualizations and the vastness of the named entities pertinent to the domain. On the other hand, efficient handling of this immensity of information requires dense conceptualisations in a highly expressive, formal ontology for it to scale well and maintain the accuracy advantage of logical inference algorithms

In order to keep user models and their usage in the information system lightweight and manageable we identify the need to build and maintain an ontological knowledge base that a) can mainly support meaningful representation of contextual (world) semantics that concern the user under *a single uniform vocabulary*, b) will be able to sustain abstract user-specific conceptualisations such as user status, skill and situation. Such an ontology should be based in a widely understandable language (e.g. English) with multilingual support, across LinkedTV scenarios. Evidently, various structured information (different sources, formats) can be integrated, mediated and exchanged with this ontology. This LinkedTV ontology can also be used as backbone for uniformly representing media and document interpretation annotation, enabling easier matching and filtering between user models and content information.

Another important issue to be clarified is the content of the ontology. Every concept, relation and rule that may have meaning to a user in the scope of LinkedTV should be represented. This means that the ontology has to be carefully designed in order to avoid non-transparent structures and ambiguity.

To this end, we will consider extending Condat's SME ontology by enhancing the English DBPedia to the three scenario languages, as will be detailed later on. Once the core knowledge is expressed in a uniform vocabulary, the NERD ontology can be incorporated in order to provide extended support with named entity recognition and Linked Media interrelations. To leverage the unmanageable size of information in DBPedia, we will extend CSME's technique to pull a relevant active ontology per user session by examining cross-topic information propagation.

However, DBPedia's expressivity is rather low. It only deals with concepts, instances and generic relations connecting them. No complex information is available conveying the distinct axioms and specific relations between concepts that adequately describe a domain. Furthermore, its size can prove to be unmanageable for intelligent inferencing services, such as reasoners, to handle. Therefore, we will consider developing three more granular use case ontologies that will further express the complex knowledge pertinent to the scenario domains and reduce domain(s) vocabulary into more compact and meaningful conceptualizations, based on identifying relevant DBPedia subsets and automatically analysing the domains.

Nevertheless, we cannot assume that domain knowledge remains static as the world changes and new trends emerge and evolve. To this end, in extension to freely available DBPedia updates, we will consider dynamically adapting the ontology to new data and evolving it over time by considering strategies to analyse aggregated user generated information to discover new knowledge about the domains.

4.3 Implicit preference mining

As aforementioned, LinkedTV aims to unobtrusively capture the preferences of the user based on his interactions and transactions within the platform. To this end, although support for explicit declaration of preferences is expected within the platform (i.e. demographic data provision, concept selection/rejection, content rating), identifying methods and sources for implicitly extracting user information is of pivotal importance to the personalisation task. Implicit user information tracking strategies should tackle the optimal trade-off between maintaining user independence (non invasive information extraction) and providing meaningful data to the profiling process.

4.3.1 Transaction tracking

Since LinkedTV functionalities will be encompassed in a unified platform, it is feasible to embed a tracking agent to the platform for more precise user logging. The tracking agent will be responsible for transactional information extraction and user logging. User transactions with web content will be tracked through a transaction listener.

To that extent, the agent will be responsible for capturing maintaining the user's viewing history information over LinkedTV content (content item id, timestamp), along with his transactional activities (skip, pause etc), as well as the interaction with external web resources. Fusion with less invasive techniques, such as cookies and web/search logs, will be considered in the case of identified multiple users or of an unknown session (no login preceded interaction with the system). In addition, the tracking system will be able to extract statistical information about the user's transaction, such as the click-through rate and the time spent on a (non-video) content item. Obviously, information such as time and location are trivial to obtain but will be decisively requested from the tracking mechanism.

In conclusion, we consider as positive transactions, hence interests, viewed content items with any complementary positive action (save, bookmark etc). On the counterpart, actions

like video fragment skipping will be denoted as a negative transaction, hence disinterest. In addition, we will consider analysing statistical transactional information, such as click-through rate or view time versus playback time and time spent on external web resources, in order to differentiate positive and negative transactions. A more detailed description of issues pertaining interest/disinterest detection will be presented in chapter 4.4.2.

4.3.2 Preference mining

The LinkedTV personalisation component needs to take into account multifaceted user behaviour for tracking user preferences. Even though the preferences themselves will be expressed in semantic conceptualizations based on a reference knowledge base, envisaged implicit preference mining cannot assume that the conceptual description of the preferences will be explicitly provided by the user. To this end, three sources of information will be considered for mining user preferences: the semantic annotation of the provided multimedia content, the conceptual liaisons of content interconnected in the Linked Media Layer and the unstructured features of web and social content, hence rendering the preference mining task a hybrid of knowledge-based, content-based and peer-based techniques. Each of these mining tasks put forward distinct requirements:

- Ability to analyse raw textual web content (tokenization, entity extraction)

- Ability to capture influence between peers and analyse peer-related content in a non-invasive manner

- Ability to produce efficient mappings between raw content descriptions (text, metadata, social media actions, e.g. likes) and predefined knowledge,

  o including ability to align the semantics of a uniform WP4 ontological vocabulary with the semantic information provided by the multimedia concepts' interconnection in the Linked media Layer (WP2)

4.4 Knowledge acquisition

Knowledge acquisition in the context of this work package refers to the basic semantic interpretation of the user behaviour, i.e. capturing the semantics of the information pertinent to the user and mapping them under the conceptualisations of the predefined ontological knowledge. Given the fact that this essentially involves understanding the user's transaction with some sort of content, whether that might be the audiovisual content that the platform provides, extra content such as articles, ads, blogs, photos provided within external Web resources or shared among peers, it is reasonable that this process mainly involves understanding what the content itself is about.

As aforementioned, although the platform's audiovisual content is expected to be semantically described and interlinked with additional information, this is not the case for freely available external content. Furthermore, the knowledge acquisition task is required to produce a lightweight and compact interpretation of the user behaviour, unifying multi-dimensional content. These challenges designate the need to fuse multilingual information,

disparate media annotation vocabularies and raw textual content descriptions under a uniform representation.

Effectively the goal of the classification task is to provide a compact semantic description about the content a user consumes and highlight the impact that this item has to the user, while simultaneously encompassing rich and deep conceptual information in this dense description.

4.4.1 Content information management

The first step in acquiring knowledge about the user based on his interaction with the multimodal content in a networked media environment such as LinkedTV involves extracting descriptive features about the content, whether that might be concepts/instances from the media annotation based on a certain vocabulary, raw text, or semi-structured metadata. Generally, the efficiency of content interpretation relies on three distinct characteristics of the input feature vector representing the content item:

- low dimensionality;

- be discriminative;

- be descriptive.

Input data

The only textual information, which comes with the content item with no machine learning effort are the audiovisual content metadata. In the case of the RBB use case, the metadata consist of a set of keywords and a textual abstract. We can assume that semantic information about the A/V content, incorporating the information encompassed in the metadata will be available following the WP2 multimedia content analysis.

Nevertheless, as the user is expected to interact with multiple heterogeneous resources, a significant number of unstructured textual information should be processed in order to adequately describe the content. We can assume that the same text mining modules used in WP2 to interpret media-related textual content can be employed to analyse the raw text in the unstructured external content.

While data in short textual descriptions such as annotations, keywords and tags alone in free-form text are low dimensional, they are not discriminative enough. On the counterpart, lengthy textual web content such as articles contains abundant descriptive information but the volume and lack of structure of their data render them neither discriminative nor low dimensional. Depending on the output format of the WP2 multimedia content analysis which has yet to be explored, some of those challenges might also be applicable to the semantic descriptions of the multimedia content, depending on the amount and particular nature of the WP2 output information.

The following is an overview of possible approaches to cope with the content features-related challenges by foreseeing a proactive analysis framework of the features in user-pertinent content and handling them in such a way that will facilitate its semantic interpretation.

Dimensionality reduction

The output of (textual) information can be used to create a Bag-Of-Words (BOW) representation of the content item. Standard text mining strategies also can provide a BOW for the analysed content. A problem which immediately arises with most machine or preference learning algorithms is the high dimensionality of the BOW representation of extensive textual information and a comparatively low number of training examples – i.e. content items viewed by the user.

Within LinkedTV, we generally envisage to use feature selection dimensionality reduction techniques. The aim is to prune the BOW representation so that only the salient words or word sequences remain. This can be achieved by entity extraction from the textual descriptions and metadata as well as from free-form text. The entities extracted comprise mainly of named entities but also of selected generic entities – noun phrases.

Ontology-based content information expansion

The extracted entities can be directly (string-to-string) mapped to URIs in the linked data cloud, and then a more robust representation might be created by using DBPedia relations to identify related instances. Related instances will be included into the feature vector with a weight taking into account the distance between the source instance and the related instance.

Information extraction-based content information expansion

Additional concept extraction can be performed by mapping the entity to its hypernym. This can be performed with Hearst-pattern extraction from suitable white-listed encyclopedic resources, particularly Wikipedia. Once the hypernym is extracted, it can be dealt with in the same manner as an entity identified in the raw textual data. The advantage of involving hypernyms in addition to the original entities is that the feature vector will be made denser.

Using the *information extraction approach* we can obtain additional entities by focusing on the hypernym relation is complementary to the *ontology-based approach* described above, which first maps the entity to DBPedia and then follows selected relation types.

Graph-based content information expansion and ontology enrichment

Linguistic relations between terms extracted from raw textual content, or even concepts of diverse semantic vocabularies, can be analysed per user, per user cluster or per domain by aggregating the information of extracted BOWs, derived from consumed or training content, into a lexical graph. The nodes of the graph contain the term (whether that might be a single word, word sequence or a URI), while its edges depict some relation between the terms (e.g. co-occurrence of the terms in the same resources) and the strength of that relationship. Other statistical or semantic information such as the term frequency (nodes) or the relation type (edges) and so on can be incorporated in the graph. As the information is gradually aggregated, the graph will be trained with relational information about the free-form data in context.

As a result, any given concept/instance in the WP4 ontology can be enriched by the propagated information derived from the neighbors of graph-related terms [Tsatsou09], where the graph neighbors of a concept (whose lexical manifestation is found in the graph) are integrated in a feature vector (per concept), these features describing the concept at hand. This information can consist of common knowledge implied in unstructured data that encyclopedic knowledge will typically discard and can provide information about the presence of a concept in a given content item even if that concept's actual lexical manifestation is not present in the item's text. Furthermore, such graphs can outline event-related information, particularly in instances of the graph of a specific time period or context.

While this method offers richer generalized information about a given resource (content item) based on common knowledge, it does not take into account structured or semi-structured world knowledge within freely available linked data. In a linked media environment such as LinkedTV it is valuable to extend graph-based analysis of free-form content in order to exploit the advantages offered by linked data interrelations underpinned within Linked Open Data (e.g. DBPedia) or in LinkedTV's Linked Media Layer, thus also alleviating the need for strenuous lexical pre-processing. In effect, graph induced linguistic relations can be extended with encyclopedic knowledge or relational information available in various disparate media vocabularies, e.g. by initiating graph construction or expanding constructed graphs with relational information available in linked data structures such as the Linked Media Layer or DBPedia or by amplifying graph edge information with linked data-based relations.

Furthermore, the ontology and information-based content expansion methods described earlier may also add to this process by e.g. identifying a more relevant concept space to constrain neighborhood (thus feature) selection for a particular concept in the graph.

The goal of this approach is to provide a method to characterise ontology concepts with relevant visual, semantic and lexical information so as to provide the basis for extracting concepts from diverse digital media content and facilitate on-the-fly recognition of the semantics in any given media resource. We will further consider projecting user-aggregated relational information to the graph, either by expanding it to a tripartite user-item-resource hypergraph or by incorporating relational information about terms in user clusters, in a post-processing step following initial single-user preference learning.

Multilingual content

While media annotations can be expressed in a "generic" language, text or audio content comes in a specific natural language. The same holds for information from social networks. As aforementioned, target users of the LinkedTV platform do not necessarily share the same language. Therefore, the user-centric ontology must map onto the user's natural language in order to be useful.

In order to avoid the tower of Babylon syndrome in LinkedTV we propose to further enrich the English-based core ontology/ies with adequate information to align multilingual information under a unified language by indexing under a single English conceptualization the corresponding information about the concepts and properties from the German, French

and Dutch DBPedia In essence, all lexical/conceptual information within DBPedia describing a concept of the reference ontology in all three scenario languages will also be a part of that concept's descriptive feature vector, while further enhancing it with synonymy information from available thesauri.

4.4.2 Semantic classification of consumed content

As aforementioned, the most prominent task in user profiling is understanding user preferences based on his transactions with multifaceted content and unifying them under compact conceptualizations. This is achieved by classifying user-consumed content to ontology concepts and/or instances.

The classification process aims to retrieve mappings between ontology concepts, instances or properties and consumed content in order to produce a dense, lightweight and meaningful semantic description of the content. It is expected that following the work of WPs 1 and 2, comprehensive mappings of multimedia content to structured ontological knowledge will be available. Nevertheless, the personalisation task is possibly expected to align mappings in several different vocabularies under a single homogeneous vocabulary (the WP4 upper ontology/ies) or be required to pursue further dimensionality reduction of media-specific semantic information in order to maintain a lightweight and agile representation schema of the user model. Foremost, a semantic user profile is required to interpret and represent user-pertinent information derived from several unstructured extra-media resources.

The challenge of interpreting and aligning heterogeneous or even unstructured content descriptions can be addressed by taking into account both available mappings of media objects to freely available ontological vocabularies as well as semantic information methodologies that rely on the linguistic analysis of textual descriptions of free-form content.

To this end, we suggest several steps in semantically recognising the multimodal digital content, commencing by classifying content items based on a broad uniform vocabulary such as DBPedia, extending the derived classification set by taking into account the expanded media information incorporated in media annotations and the Linked Media Layer and then reducing it to more compact user-specific semantics based on information encompassed in a pre-trained lexical graph that can take into account both lexical, encyclopaedic and linked data information, such as the one described in chapter 4.4.1.

Partner Condat's SME semantic fingerprinting methodology can serve towards mapping content to DBPedia. In the CSME extracted entities from a content resource are mapped to URIs in the linked data cloud with DBPedia being the primary resource. Such a semantic fingerprint can be extended and aligned with the semantic annotation of multimedia content foreseen within LinkedTV, supported by available mappings to DBPedia and other descriptive semantic knowledge bases and interrelations of media content within the Linked Media Layer.

Nevertheless, the need to reduce this information stems from the fact that classifying content to such a broad concept space places the classification load directly on the server, due to

pedestals in performance posed by the volume of its employed knowledge bases, while desisting from making the most of the formal semantics incorporated in domain-specific ontologies. Nevertheless, it can still provide seed information to instigate content-specific knowledge discovery, on the server or a mediation layer, which can be imported to the client for more targeted and advanced classification of user behaviour.

Hence, we can employ lexically enriched (as described in chapter 4.4.1) domain-specific ontologies, where concepts are characterized by the lexical and semantic information encompassed in their descriptive features, e.g. multilingual alternates and related graph neighbours. In essence the goal of this approach is to provide a bridge between the shallowness, diversity and highly dimensional of linked data information in vocabularies such as DBPedia and the Linked Media Layer and the narrower and more granular conceptualizations of the use case ontologies.

Each extracted term[25] in the text body or annotation of a content item, whether that might be a multimedia resource or a text resource, can be looked up in all the feature vectors of the concepts in an enriched use case ontology, in order to retrieve ontology concepts that this term semantically characterises. We should note that a given term might characterise more than one concept – most probably with a different degree, that degree denoting the strength of the relation of the term to the concept. Consequently, a set of concepts and/or instances from the reference ontology is retrieved along with a standing weight, that weight denoting the participation of the concept to the content. These concepts/instances constitute the classification set $\{c_1 \cdot w_1', a : c_2 \cdot w_2', ..., c_n \cdot w_n'\}$ for the particular content item, where c is an ontology concept, $a{:}c$ is an ontology instance and w in the assigned degree of confidence that the entity represents the content item.

We will further investigate the opportunity to exploit CSME's mappings to DBPedia and the Linked Media Layer semantics in order to extract specific properties quantifying classified concepts/instances.

4.5 Profile learning

The conceptual representation of atomic (single-concept) user preferences based on the uniform WP4 upper ontology allow us to update and extend the user profile in two distinct modes: assign, update and revise the degree of interest/disinterest a person has for a given concept and discover complex relations pertaining mined concepts that illustrate persistent behavioural patterns of the user.

---

[25] As a term here we denote the lexical manifestation of any given textual token, whether that might be a word, a sequence of words or a concept/URI.

4.5.1 Stereotypes

A very initial insight about user preferences can be achieved by categorizing the user under several stereotypes based on his (explicit) personal information, without him even having to have interacted with content or peers within the platform. For instance, the usage of social media networks such as Flickr, YouTube, Facebook, LinkedIn for data mining issues provides an exceptional quality of explicit user information. Based on such information stereotypes can be detected by various clustering methods.

Recent advancements in the development of ontologies for user profiles as described in section 2.4.1 and 2.4.4 open opportunities for formalizations of single stereotype models as separate instances of the ontology. An implementation of a stereotype as an instance ontology makes it a sharable and reusable unit and can lead to the creation of libraries of such modularized stereotypes. In addition, this method can serve as a fist-level knowledge pulling mechanism, providing only the segment of the ontology that is relevant to the user, hence achieving reduction of the dimensionality of available reference knowledge and supporting more lightweight data handling.

An adaptive system relying on stereotype-based modelling does not update every single facet of the user model directly. Instead, it utilizes a stock of preset stereotype profiles. Whenever the system receives a hint about a user being characterized by a certain stereotype, the entire user model is updated with the information from that stereotype profile. A user can be described by one stereotype or a combination of several orthogonal stereotypes.

4.5.2 Interests, disinterests and their importance to the user

User preferences are not limited to interests. Disinterests are perhaps even more substantial to the personalisation process, since avoiding recommending content that the user is indifferent to or even strongly against can substantially enhance recommendation quality. As mentioned in section 4.3.1, several types of tracking information can aid the profiling process to discern between interests and disinterests, the most prominent being the actions the user performs on the platform in interaction with the content as well as several other aspects like statistical analysis of transactional behaviour.

Therefore the primary step in effectively learning user behaviour is yet another classification task which is far from trivial: discerning between interests and disinterests. An interactive TV context, offers extended information on the nature of the user involvement with the content than generic information systems due to the more expressive modalities and facts offered by the video reproduction interface and the nature of the audiovisual content. Hence, a video player offers the opportunity to track a finite set of user actions concerning the content such as pause, rewind, skip etc that can bare a positive or negative impact to the user, further enhanced within LinkedTV by the provided ability to interact (or not) with triggerable video objects. In addition, the fact that the media content (or media fragments) carry a particular transaction timeframe (duration) that binds user interaction with the content, we can

statistically determine the interest probability of the user based on time spent on the content item (that being the entire video or a specific fragment).

Following identification of whether a content item is interesting or not, the system proceeds to understand the content by semantically classifying it to available ontological conceptualisations as illustrated in the previous chapter, potentially providing a classification confidence degree. Identified disinterests can be semantically interpreted as the complement of the profile, i.e. all that is not the profile - in section 4.6, we will illustrate how this interpretation can be semantically interwoven with interests.

Consequently, the weight of individual preferences can also be modified by semantic weight modifiers [Bobillo08] in order to convey the impact of the preference of that concept to the user. Weight modifiers can either be predefined based on the significance of user actions (bookmark, skip etc) or statistically elicited by the user's attention span, effectively reflected in interaction manifestations such as click-through or time spent on a video content or page – or both.

As the user consumes content over time, simple semantic preferences will be aggregated to the profile (i.e. atomic or property-restricted concepts/instances). Nevertheless, significant shifts in long-term user interests are expected to occur. Two factors affect a preference's impact over time: the frequency with which the concept appears in consumed content and the age of the concept in the user profile. To this end, the preference weight will be continuously adapted based on concept frequency and a time decay factor. In order to illustrate a preference's impact to the user in proportion to the entirety of his preferences and facilitate preference ranking, we will consider using a finite weight range (e.g. possibly $w \in [0,1]$).

Furthermore, since long-term interaction with the platform can lead to an unmanageable number of user preferences and an unobtrusive personalisation system cannot expect from the user to explicitly capitalize on his ability to delete, modify and flush his profile, the preference weight can be used for automatically ranking and pruning preferences. As a result, the profile adapts to fluctuating user preferences, with new concepts incrementally appended, and/or eliminated, or weight-modified.

4.5.3 Object ranking preference learning

After obtaining a set of (weighted) atomic user preferences based on the aforementioned methodologies, the system will be able to re-assess and adjust the interest weights based on the utility of the content with regard to the user. The input for such an object ranking profile learning method is constituted by a semantic and topical representation of content items viewed by an individual user, user interest level (weight) and the context.

The *semantic representation* of content items amounts to a weighted vector of concepts from the content ontology. In this manner the semantics of the content item can be reasonably precisely expressed, however at the cost of higher dimensionality of the representation.

In contrast, to the semantic interpretation the *topical representation* intends to give another - a more high level - view of the content items. It can be roughly compared to the rankings given in TV guides in the form of a number of stars in several widely acknowledged criteria e.g. (suspense, romance, etc). The topical representation can either be sourced from meta-data provided by the broadcaster or generated automatically by text categorization algorithms. This representation is low-dimensional.

User interest level is computed from data on user interaction with the content item and the assessment of user attention as judged from input from external devices.

UTA Method

In the LinkedTV framework, it is intended to use the explicit and particularly the implicit feedback from an individual user to sort the individual content items the user interacted with according to the degree of her interest. Such a representation of user preference – alternatives sorted from the most preferred alternative to the least preferred one – corresponds to the stated order of alternatives, as known from the field of Multicriteria Decision Analysis [Siskos05]. One of the most suitable approaches to evaluating data of this kind is constituted by utility based methods [Furnkranz10]. Of particular interest is the UTA method, which is a disaggregation-aggregation method for preference learning and rule-based methods.

As was pointed out in subsection 2.2.2., the UTA method seems suitable for LinkedTV since it encompasses two important advantages compared to most other preference learning methods.

Firstly, the output of the method is a model explaining user preferences in the form of utility functions, which is perhaps the most widely known representation of preferences. In this respect, UTA method best matches the requirement that the user should be able to inspect her preference model. Utility curves learnt by the UTA method can be easily visualized and prospectively even edited.

Second, since the UTA method has a very strong inductive bias, the number of required training examples is smaller than for other methods. This is a significant advantage given the fact that the user starts with no training examples.

UTA method draws inspiration from the way human handle decision making tasks at a level of abstraction provided by the microeconomic utility theory. The knowledge of the way humans tend to make their decisions can serve as inductive bias in a machine learning algorithm. Inductive bias of a learning algorithm is a set of assumptions made about the target function. The more valid assumptions are incorporated into the learner, the better the approximation performance on unseen data keeping the same number of training examples. The principal inductive bias to be used in utility-based methods is the monotonicity of utility functions. The UTA method also incorporates additional assumptions, particularly piece-wise linearity and additivity of utility functions. Example of three piece-wise linear monotonic utility functions are below. In the example, u_1 could correspond e.g. to "romance", u_2 to "action" and u_3 to "politics".
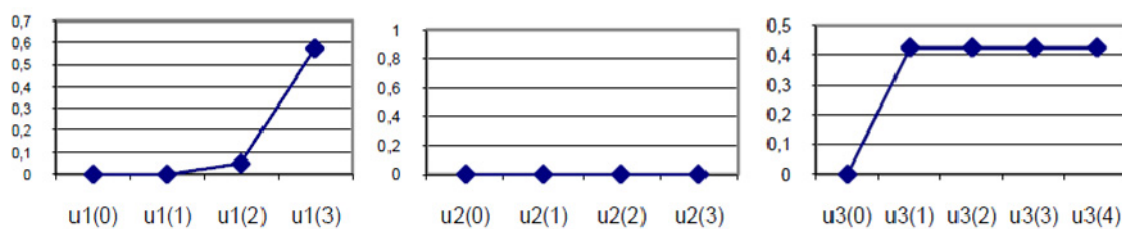
Figure 7: Piece-wise linear monotonic utility functions

In a testing phase, unseen content items can be assigned a utility according to the utility curves in individual criteria, the utility values are summed and the content items sorted according to the sum.

One of the requirements of the UTA method is the ordinal nature of the input criteria. While the assumptions relating to the monotonicity of preferences have been partly addressed in our previous work [Kliegr09], the performance of the method would fundamentally suffer if there are many nominal criteria. Also UTA method has so far been typically applied in domains with relatively few criteria. The comprehensive visualization benefits would be also lost if the number of utility functions to be presented to the user is too high.

It follows that the benefits of the UTA method will be best utilized if applied on the low-dimensional topical representation of content items. In this representation, there is only a small number of ordinal criteria. While it cannot be assumed that the criteria are fully monotonic with respect to the user's utility function, the assumption that there is a maximum one peak seems to be acceptable. A limited non-monotonicity of a criterion can be coped with by our non-monotonic extension of the UTA method [Kliegr09]. Using the TV guide analogy, the user's preferred value for the "romance" criterion may be "***", rather than no romance at all "*" or maximum romance "*****".

It should be noted that the UTA method will be taken as a starting point for utility-based learning within the project. Based on the research carried out within the project, the method will be further developed or possibly superseded by another utility-based algorithm.

4.5.4 Rule learning over ontology-based user profiles

Rule learning is expected to enable the opportunity to highlight patterns (i.e. complex, non-atomic user preferences) in user behaviour and represent them in a machine-understandable format. Even when run on high dimensional data, individual rules tend to be short, which fosters human readability of the rules. It should be noted that depending on the mining setting and selected algorithm, the result of the rule mining mechanism can consist of a relatively large number of rules. For example, in [Dembczynski10] two algorithms - RankRules and PrefRules - were applied on the same problem. While the RankRules algorithm required 100 rules, the PrefRules algorithm gave the best performance with only 10 rules.

The choice of the rule learner

The number of rules produced depends on the conciseness of the rule formalism on the output of a rule learner. To this end, it seems desirable that the rule learner can produce rules with the full range of logical connectives, including disjunction and particularly negation.

In context of the LinkedTV project, rule learning will be applied in highly dimensional attribute space. The availability of the underlying ontology, gives the possibility to use a taxonomic relation between attribute-values also during the mining process.

These requirements are fulfilled by the GUHA-method rule learning algorithms (also refer to 2.2.4). The LISp-Miner (Simunek)[26] implementation of the GUHA method ASSOC is a rule learner with support the required range of logical connectives - conjunction, disjunction and particularly negation, which can be e.g. used to express disinterest of a user.

Harnessing the ontology

The possibilities of using ontologies in data preprocessing in association rule mining were investigated e.g. in [Svátek05] and for postprocessing in [Marinica10]. The problem with the former is that it is not clear to what granularity it is best to preprocess data before mining. A partial solution is to use multiple granularities, however at the expense of increasing the size of hypothesis space. While applying ontologies to prune discovered hypotheses in the postprocessing step may be effective, it requires the complete hypothesis space to be previously mined, which may not be computationally feasible.

The LISp-Miner system offers a feature called classes of equivalence, which gives a limited possibility to cope with the ontological structure of data on the level of taxonomic relations during mining. Apart from producing shorter rules without loss of semantics, the benefit is a possibly significant decrease in the size of the hypothesis space. This is particularly important, because dimensionality of the semantic representation to be mined is potentially considerably high.

LISp-Miner is a grid-enabled software. Within the scope of the LinkedTV project, it is planned to strengthen the computation capacity of the UEP hosted grid, which currently reaches only proof-of-concept levels.

The technological challenges related to rule mining include apart from the large dimensionality of the input data also dealing with the evolving nature of the preference dataset, as new instances appear as the user consumes additional content and oldest instances are removed.

4.6 Profile representation schema

The decision about the most appropriate user profile schema to be adopted within the LinkedTV framework depends entirely on one pivotal notion: using lightweight user

---

[26] lispminer.vse.cz

knowledge structures that would enable us to efficiently manage (user and reference) knowledge and make meaningful inferences even in limited resource devices in order to reduce server communication load and ensure user security safeguarding. To this end, we refrain from employing mere feature vectors of weighted concepts/instances that accumulate user preference in order to be able to depict complex user-specific knowledge such as rules and to efficiently handle disinterests in parallel with user interests. Similarly, with an additional interest towards handling scalability issues with respect to both faster inferencing mechanisms and towards preserving user privacy by storing the profile on the client only, we abstain from using an activated instance of the ontology for profiling the user.

Instead we proposed an axiomatic representation of the profile, within the DL expressivity fragment, such that the union of all interests comprise the profile and the union of all disinterests comprise all that is not the profile:

$$\bigcup_n \exists hasInterest.Interest_{i \in n} \rightarrow Profile \, ,$$

$$\bigcup_n \exists hasInterest.Di\sin terest_{i \in n} \rightarrow \neg Profile \, ,$$

such that $Profile \bigcap \neg Profile \subseteq \perp$.

The latter essentially denotes that given that an entailment exists for the profile, if an entailment also exists for something that is explicitly not interesting to the profile (disinterests), the hypotheses leads to a refutation, therefore the profile criteria cannot be fulfilled.

An interest can be represented by an atomic concept, an instance, or a complex concept, i.e. a set of concepts related through some constructor, i.e. conjunction, disjunction or negation (e.g. from mined rules). Concepts and instances can also be existentially or universally quantified to ontology properties.

We will consider a highly expressive representation language that can cover DL semantics such as OWL or any other language that can be employed within the expressivity fragment of the LinkedTV scope in order to capitalise on its capability to express knowledge in formal ontologies and support for essential semantics such as relationships between classes (e.g. conjunction, disjunction, negation, disjointness etc) and richer properties' semantics. The popularity of DLs issues not only from the direct relation with OWL, but also from the fact that they constitute expressive fragments of first order logic, for which decidable reasoning algorithms exist, widely used for, but not limited to, semantic representation of multimedia content [Dalakleidi11], thus providing an optimal framework to unify multimedia and generic content semantics

Semantic interpretation of preference rules

In order to semantically interpret the aforementioned mined rules to a machine-understandable format suitable for inferencing, a semantic axiom will be created and updated per rule for the user. It is evident by aforementioned specifications that a simple rule will contain two parts, one being the conditional relations between ontology entities (e.g.

negation, conjunction, disjunction) that make up the rule and the other concerning the conditional impact of the rule.

Therefore, a primitive naïve interpretation of such a rule would be translated into an implication axiom, where the body contains the complex relationship and the head would contain an interpretation of the rule's impact by attributing an impact weight to the rule.

As a naive example, given a simple rule: *football (high) and berlin (low) → 0.9, attention_level(low),* where → *0.9* denotes the strength of the rule with respect to these measures and the rule reading "*if football is high and Berlin is low then in 90% of cases the user attention level was low*", refraining from defining the statistical intervals inside which the constants "high" and "low" are defined, an axiomatized interpretation of the rule would examine the following parameters:

- Is the rule positive or negative? → *attention_level(low)* → negative

- What's the strength of the rule? → 90% 'hit' rating. → Strong, with a strength of 0.9

The primitive axiomatic interpretation of the rule could reflect an implication rule where the >0.7 · Football ∩ <0.4·Berlin ⊆ Rule1*0.9. The rule-concept would be inserted as a complex concept to the user disinterests. Further research will aim to more substantially express the impact of the rule to the profile based on the diverse statistical information available from the elected rule learning algorithms.

4.7 Storage and communication

The storage and server-side communication of personal user data is a most crucial issue for LinkedTV's personalisation task. Particular interest will be given in order to ensure the security of sensitive user information. Hence stems an overarching interest of this work package: to focus on encompassing user preferences in conceptually dense but lightweight models that will enable their storage and usage directly on the client for semantic inferencing, i.e. for content/concept recommendation.

At an initial phase, we will deploy anticipated profiling techniques on the server for research purposes, considering however future actions that will be needed in order to minimize client-server communication and secure user privacy wherever communication to and from the server is needed. Nevertheless, even initial planning of the foreseen client-server architecture ensures that no individual user personal information is required to be stored on the server. Rather, we will highlight which user preference knowledge acquisition and learning methodologies require the extended processing capabilities of the server and indentify a strategy according to which communicated data will be anonymised and encrypted, as well as removed instantly after they are processed and their output is transmitted back to the client.

To this end, we anticipate that a transaction listener will be embedded in the Linked platform and information tracking will be performed on the client. The ontology/ies however will reside on the server and cannot be transported to the client in full length; therefore a mediation

cloud will be considered where an active ontology will be pulled for classification based on contextual information, potentially combined with user cluster-based knowledge (e.g. particular events, group-pertinent complex knowledge etc).

Concerning the latter case it is anticipated that new group knowledge might arise and that it will need to be stored and propagated in the initial reference ontology/ies. In this case, the user-contributed group information that is communicated back to the server will be aggregated and anonymised and no identification details about an individual user of the cluster, encoded or otherwise, will be imparted.

We will also consider possibilities of extending and combining several more knowledge pulling techniques in conjunction with the use of a lightweight and expressive domain-specific knowledge, as already considered, in order to potentially enable the integration of the classification task on the client.

In any case, in future architectures the individual user profile is foreseen be stored and be made available for inferencing directly on the client, while appropriate inferencing mechanisms that can perform efficiently on the client will also be considered. The main goal is to render the proposed personalisation method agile and adaptable while restraining server load and maintaining user privacy.

5 Contextualizing user behaviour: requirements & specifications

The concrete context of a user has, of course, a great influence on his decisions about media consumption. The user is influenced in various ways by her social environment (e.g. family, friends and social networks; as well as news, media, and big events). Also his concrete situation is influential: the time of the day (she prefers listening to the radio in the morning but likes art house movies in the evening), time of the year (Christmas, summer, holidays, etc.), from the location (at home, at work, at a business travel, on a holiday trip), etc. More importantly, the physical state (alone, with company) and reaction to the content (mood, attention) are prominent indicators of the contextual situation of a user.
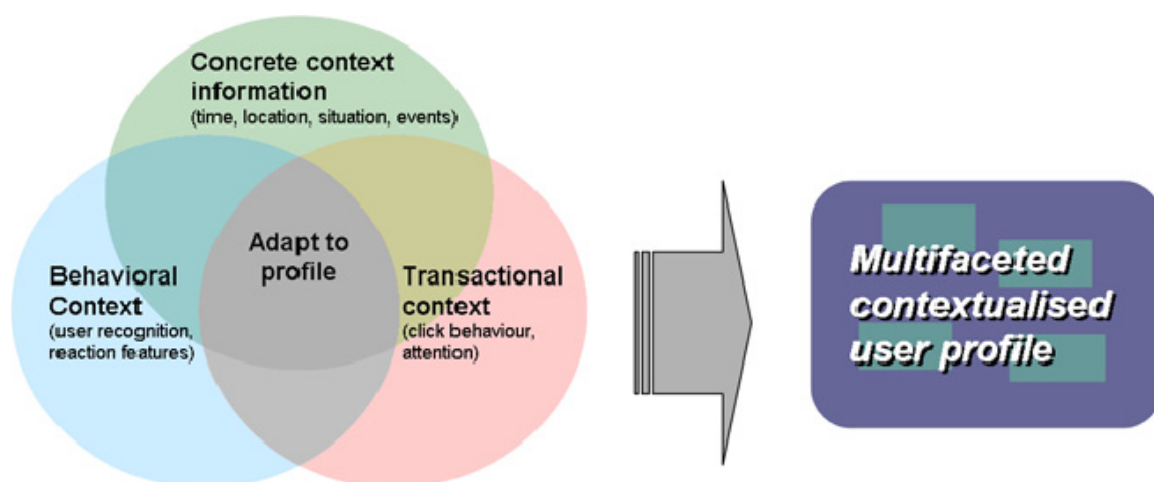
Figure 8: Context information

5.1 Functional requirements

A list of user-related features (reactional[27] & transactional) that will describe his behaviour will be extracted and used to determine the user's situation. Depending on the scenario, the list of the extracted features can evolve. Concerning the physical (reactional) behavioural tracking, there exists a finite list of all the features that can be extracted from user behavior which will be validated by the ethics committee, but if a given scenario does not require the entire set of features, the corresponding list will be limited to save computational power and cope with privacy protection.

A typical scenario where news is displayed depending on the user profile would simply require:

---

[27] Refer to section 5.2

- the number of present persons to adapt to one or several user profiles,

- user recognition to know which profile to attach to him and probably

- motion (excitement), audio and focus features to know if he reacts to the content and when exactly he reacts for further content annotations.

More complex scenarios as TV games or interactive artistic performances on TV would need more fine people motion analysis (symmetry, contraction index) and also interpersonal distances and their variations. Each scenario is required to provide a description file where the needed features are described.

5.2 Tracking user behaviour in reaction to the content

Here, we will extend the [Ballendat10] approach based on Greensberg's 5 features [Greensberg11]. We plan to use a Kinect sensor which is low-cost and common. This sensor can be easily plugged in a set-up box by using USB plugs and it is already widely accepted by a large public which is a good point from a psychological point of view. From the Kinect sensor RGB, depth and audio features can be extracted.

Based on the [Greensberg11] approach we will have an extended five points approach:

- Location: the scene will be automatically scanned by the Kinect sensor and a 3D environment map reconstructed. This 3D map will provide the system with information about the environment close to the TV. During time, the system will be able to learn where people come from and where do they leave (in/out regions) but also where people have a high probability to focus on the TV (sofa) or to talk together (dinner table). Finally the system will be able to see how people interact with objects surrounding them. For this purpose several steps are necessary from visual SLAM techniques (3D reconstruction) to behavioral learning (when people sit there, they are likely to focus on the TV).

- Identity and context: knowledge about the identity of a person. A person can be recognized by using several features like speech, biometric information or face information. We will test those which need less learning. A combination of face recognition and biometric features (tall, small, fat, gender…) seems to be a good choice. This point is also important to know the number of people, if they are already known or not, and to extract biometric features mainly about their age and gender.

- Orientation: the relative angles between entities. It is possible to extract the normal direction of the body so that we can extract people orientations relative to other people or relative to environment objects (extracted during the location step). This point is important to detect if the focus of attention is on the TV or not for example. Audio features will also be used to see if there might be a conversation between people, or if someone is talking.

- Distance and static features: the distance between people (interpersonal distances) and objects of the environment can be extracted from users' and objects positions.

The distance of the users relative to the TV can be used to activate implicit or explicit interaction or to understand the relations between the different users and the TV (who is really interested, who is just there to talk to the others…). Other static features will be extracted from people silhouettes like their symmetry (how symmetric are the hands, is there one hand pointing the TV screen …), the contraction index (are the hands along the body or not). Figure 9 shows the environment reconstruction and the user detection along with his interpersonal distances (intimate space in red and social space in blue).
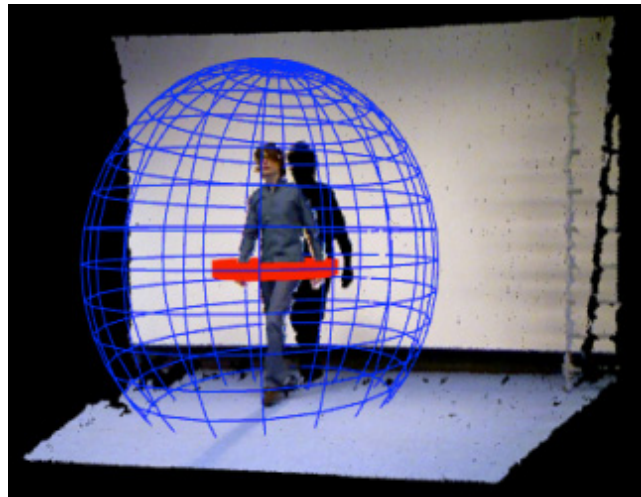
Figure 9: User detection and interpersonal distances display

- Motion and dynamic features: changes of distance and orientation over time are interesting to analyze the evolution of the interest of people in the content delivered by the TV. The motion index (which is the global motion of the body including hands, feet shakes) will also be taken into account. Finally, audio features can also provide cues about the level of excitement and emotion of the user. It is also crucial to know at which moment an important change in the previous behavior occurs. This can be due to people arrival, remembering of a task which should be completed, the TV content does not fit anymore or on the contrary fits much better with the user needs. The motion features can help in context change detection, which is very important in changing the user profile. Figure 10 shows dynamic features extraction from users.

Figure 10: Dynamic features extraction

The behavioral tracking of user reactions to the TV content will provide with knowledge about:

- People identity (to be linked to the profile)

- People positions, biometrics (to be linked to the context)

- People focus of attention, emotional state (to be linked to data annotation and to the profile)

- People abrupt change in behavior (to be linked to possible context change and data annotation)

The list of basic features of this section will be formalized and proposed to the ethical committee for validation. Depending on the scenario all or a subset of the features will be needed. If a feature subset is needed, this will save computational resources.

The behavioural tracking can be performed for a person alone which is the simplest situation or with several people. When several people are present in the room, either one person only is taken into account (the one sitting on the sofa, or the one whose precise profile is used for example). There is also the possibility to use the mean features extracted from several people if they all look to their TV. In a first step, one viewer alone will be taken into account, while in the following tests several ways of handling multiple viewers will be investigated.

5.3 Determining user behaviour

The contextualisation approach followed in LinkedTV will resemble the principles of the CSME's semantic fingerprinting technique in the sense of producing multiple contextualised fingerprints regarding the user's situation based on the preferences in his long-term profile and related context-defining features. Besides the physical features described in section 5.2, transactional features might include temporal (time of day, season etc) or location information or the specific sub-topic that the user is currently or recently engaged in.

In particular, for every individual user there is a user model that contains the entire information needed for characterizing the user's long-term user preferences. The adaptive system manages the user model and generates different contextual fingerprints for each context determined. Each contextualised fingerprint can contain several predetermined characteristics describing the user's behavioural, transactional and concrete situation and hence is able to identify a specific peculiarity of these characteristics.

For example, from past events the system has learned that in 90% of all observed sessions at home between 8pm and 11pm the user prefers to watch the daily news (politics, economy, no sports/weather) and action movies with American actors when he is alone, expressing different disposition and attention to each conceptual subject. Accordingly, a corresponding fingerprint could look like this:

> *Time*: 8 pm – 11pm, daily (without weekend)
>
> *Location*: at Home
>
> *Concrete media consumption*: news (politics, economy); movies (action); celebrities (American actors)
>
> *Situation*: private activities
>
> *External events*: daily news, breaking news (politics, economy), electronic program guide (new action movie available)
>
> *Influences from social networks*: YouTube video about specific American actor available
>
> *Reaction*: news: attentive, movies: emotional

The information in each individual fingerprint is then semantically interpreted, by recognizing which concepts and rules in the user profile are pertinent to the context. A subset of the long-term profile is subtracted and stored separately, along with the weight that denotes the strength of the context for the user, used for pruning contexts in which interest has faded. The representation scheme of the context model will follow the lightweight representation schema of the long-term user model. In that way, the recommendation engine can only use that particular, less dimensional, subset of the profile.

It is conceivable that the described learning algorithms will be applied for each individual context model as well in order to elicit the particularities of the context in question, i.e. it could be found that the strength of a concept is much higher for the context at hand than in general or that different rules may pertain the concepts within the particular context.

5.4 Storage and communication

The contextualisation process will follow the same privacy-preserving principles as the personalisation process (as described in chapter 4.7), i.e. lightweight CUMs in the same representation schema as the long-term user profile. We will consider strategies where the CUMs can be stored on the client (for industrial purposes) while exploring methodologies to handle required communication with the server for processing purposes with attention towards proper anonymisation and encryption of user data.

Concerning reactive behavioural tracking specifically, where very distinct and sensitive user information about the user is utilised, future work will be particularly attentive towards securing user privacy. The potentials and constraints of minimising user data transmission to the server have already been initially considered. We recognize that the client storage capacities will be limited and devoted mainly to movies or other TV shows recording. Thus, the analysis storage capacities will be very small. Therefore, only the small list of features defined in section 5.2 will be extracted and a history of those features will only be stored on the disk of the client. No recording of data coming from user behavior will be performed (neither audio, RGB nor depth maps). This history will be used for example for context change detection and it will be a mid-term temporary history. Long-term data can be extracted again from those features and expand the information of the long-term user profile.

The rough (camera-extracted) data will pass two analysis filters before being stored at long term on the disk: a feature extraction which is a massive summarization of the data and then, a second analysis step will follow where new data will be integrated into the profile. The features which are extracted from the user behavior will be validated by the ethics committee.

From this data stored on the client system, a third filter will require transmission of user data linked to the contextualized profile partitions to the server. Data linked to the profile can be sent to the server in order to provide information about what parts of the TV content was considered as important based on users' reactive behaviour to help the annotation process. This data communicated to the server however will be anonymized without a specific connection to individual user profiles but can rather be used as aggregated information. But most of the communication will be incorporated in a download stream as the client will elicit the personalised content to be delivered and just sends to the server the delivery request. This content will be thus sent back according to the user scenario.

As mentioned before, securing sensitive personal user information is a crucial issue to ensure acceptance of the system by the users, so a lot of effort will be focused towards avoiding any violation of user privacy, while considering that:

- Absolutely no rough (camera-captured) data will be permanently stored anywhere (client or server).

- The features extracted from the rough data will be stored in a temporary location and only on the client. After these features' are semantically interpreted and their contextual impact is conveyed back to the one or some of the contextualized user profile partitions, they will be permanently deleted. The maximum list of extracted features will be validated by the ethics committee and this list can be shorter depending on the scenario.

- Privacy-preserving efforts will be directed towards determining an appropriate workflow concerning storage and management of the contextualized user model for commercial usage purposes.

6 Conclusions

Prior and current work on targeted information delivery systems in networked media environments indicates that two main issues arise in capturing and interpreting user behaviour for personalizing and contextualizing a viewer's experience: the lack of sufficient descriptions and metadata so as to provide substantial interpretations of viewed multimedia content and the difference in vocabularies used to describe multimedia and textual content pertinent to a user.

While within the scope of LinkedTV the former challenge is expected to be addressed in WPs 1 & 2, an advanced personalisation system can only capitalize from the information available in user-consumed data by understanding and unifying digital media data representations from various heterogeneous sources such as video, audio, text and social media, based on the user's transactional and reactional behaviour with such content. Appropriate co-operation with WP2 will be ensued in order to manage the workflow between WPs 2 and 4 with regards to interpreting user-consumed content.

Nevertheless, in order to cope with the diversity and discrepancy of user-pertinent information in such a multimodal networked media environment, while being able to efficiently utilize this information for intelligent content and concept filtering, an overarching requirement emerges: the need to underpin user preferences in a uniform, lightweight model that however encompasses a deep and meaningful understanding of user-related information. To this end, a networked media semantic personalisation system needs to:

- Understand what the data pertinent to a user denotes, based on his transaction with and reaction to diverse content resources, including:
  - o Unobtrusively extract information about the user and evaluate their nature (interest vs. disinterest)
  - o Analyse that information and acquire meaningful semantic knowledge about the user.
  - o Align disparate data stemming from content consumption, interactions in social networks, predefined domain knowledge, and explicitly defined user preferences.
- Put acquired knowledge down in a uniform machine-understandable vocabulary, considering:
  - o Using and maintaining expressive and lightweight ontological knowledge base/s
  - o Producing and maintaining a compact and meaningful semantic (ontology-based) user model
- Learn user behaviour by:
  - o Estimating preference impact
  - o Discovering relations between preferences (behavioural patterns)

- Harvest mass intelligence via:

    o User clustering

    o Stereotypes

- Contextualise user preferences:

    o Determine behavioural (reactional), transactional and physical situation of the user

    o Understand it and adapt the user profile by defining multiple contextualised facets of the user model.

This deliverable has provided an overview of the state-of-the-art highlighting how these issues are addressed in the literature, while focusing on the specific aspects of personalisation and contextualisation within networked media environments to pinpoint the main challenges arising in LinkedTV. We have outlined the main requirements pertinent to the LinkedTV scenarios and user needs and specified a set of comprehensive goals and a workflow for the personalisation and contextualisation tasks in LinkedTV, primarily based on the user's behaviour with regard to the content.

Following the identification and initial consideration of how to address these fundamental challenges in user profiling and contextualisation, we will further explore and foster manifold user knowledge acquisition and adaption techniques, such as knowledge pulling and alignment, to facilitate intelligent and privacy-preserving concept and digital content recommendation. Apart from content-specific information extraction techniques, future work will also take into account the factorisation of peer-based user information stemming from social media interactions. Additionally, we will investigate methods to expand the platform's capabilities by extending reference (ontological) knowledge through automatically learning group-specific knowledge based on aggregated user information and adapting new information to the initial knowledge.

In the interim, an imperative challenge of the personalisation and contextualisation task is ensuring the protection of user privacy, both in regard to already designed approaches as well as to future work. To this end, while initially implementing our approaches on the server for research purposes, we will further consider appropriate techniques in order to reduce the dimensionality of contextualised user profiles and the complexity of some processing techniques, thus osculating the possibility of minimizing server-client transmission of sensitive data. In addition, particular attention will be paid towards identifying and applying suitable encryption and anonymisation techniques in order to secure transmission of personal data. The purpose of this effort lies within the requirements of rendering our system competitive and suitable for commercial exploitation. To this end, corresponding actions towards achieving this goal will be cleared with the ethics committee.

The privacy preservation task will subsequently involve close interaction with WP5 in order to design an appropriate client-server workflow, suitable for the LinkedTV platform. Apart from the collaboration with WP5 on the architectural design and the aforementioned liaison with

WP2, WP4 will also co-operate with WPs 2 and 6 in terms of determining the most suitable uniform vocabulary (ontology/ies). This vocabulary will enable us to semantically represent the information provided by the Linked Media Layer and the multimedia annotations, as well as address the requirements of the LinkedTV scenarios. Finally, requirements and possibilities posed by the design of the user interface (WP3) will be factored in the user behaviour parameters to be considered within WP4.

Ensuing work will focus on implementing the methodologies described in this deliverable, while a technical overview of the deployed work on personalisation (up to month 12) will be released with D4.2. In conjunction, after determining the primary factors that will formulate the contextualised user profile (i.e. capturing and learning methodologies, schema, ontologies) we will endeavour on determining and prototypically implementing methodologies to exploit produced CUMs in order to provide targeted recommendations of concepts and/or additional content to the user of the LinkedTV platform. D4.3 will ultimately illustrate our first decisions towards producing such a concept/content filter.

7 Bibliography

[Koch00]   Nora Koch. Software Engineering for Adaptive Hypermedia Systems. PhD thesis, Ludwig-Maximilians-University Munich/Germany, 2000.

[Fürnkranz10]   Johannes Fürnkranz and Eyke Hüllermeier. Preference Learning: An Introduction. In Preference Learning (Johannes Fürnkranz and Eyke Hüllermeier, editors), pages 1-17, Springer-Verlag, 2010.

[Karatzoglou10]   Alexandros Karatzoglou, Xavier Amatriain, Linas Baltrunas, and Nuria Oliver. 2010. Multiverse recommendation: n-dimensional tensor factorization for context-aware collaborative filtering. In Proceedings of the fourth ACM conference on Recommender systems (RecSys '10). ACM, New York, NY, USA, 79-86. DOI=10.1145/1864708.1864727

[Desarkar10]   Maunendra Sankar Desarkar, Sudeshna Sarkar, and Pabitra Mitra. 2010. Aggregating preference graphs for collaborative rating prediction. In Proceedings of the fourth ACM conference on Recommender systems (RecSys '10). ACM, New York, NY, USA, 21-28. DOI=10.1145/1864708.1864716

[Ma08]   Hao Ma, Haixuan Yang, Michael R. Lyu, and Irwin King. 2008. SoRec: social recommendation using probabilistic matrix factorization. In Proceedings of the 17th ACM conference on Information and knowledge management (CIKM '08). ACM, New York, NY, USA, 931-940. DOI=10.1145/1458082.1458205

[Koren09]   Yehuda Koren, Robert Bell, and Chris Volinsky. 2009. Matrix Factorization Techniques for Recommender Systems. Computer 42, 8 (August 2009), 30-37. DOI=10.1109/MC.2009.263

[Mairal10]   Julien Mairal, Francis Bach, Jean Ponce, and Guillermo Sapiro. 2010. Online Learning for Matrix Factorization and Sparse Coding. J. Mach. Learn. Res. 11 (March 2010), 19-60.

[Jambor10]   Tamas Jambor and Jun Wang. 2010. Optimizing multiple objectives in collaborative filtering. In Proceedings of the fourth ACM conference on Recommender systems (RecSys '10). ACM, New York, NY, USA, 55-62. DOI=10.1145/1864708.1864723

[Mobasher07]   Mobasher, B., 2007. Data mining for web personalisation. In The Adaptive Web: Methods and Strategies of Web Personalisation, Springer-Verlag, Heidelberg, Germany.

[Liu10]   Liu, X. & Murata, T., 2010. Detecting Communities in Tripartite Hypergraphs. Communities, p.4. http://arxiv.org/abs/1011.1043.

[Newman03]   Newman, M.E.J., 2003. The Structure and Function of Complex Networks. SIAM Review, 45(2) (pp 167-256)

[Rich79]   Elaine Rich: User Modeling via Stereotypes. Cognitive science 3, pp. 329-354, 1979

[Kay94]   Kay, J., 1994. Lies, damned lies and stereotypes: pragmatic approximations of users. Proceedings of the 4th Int.Conf. on User Modelling (UM'1994), Hyannis, MA, USA, pp.175–184.

[Chin89]   Chin, D., 1989. KNOME: Modelling What the User Knows in UC, User Models in Dialog Systems, Springer, Heidelberg.

[Gawinecki05]   Gawinecki, M., Kruszyk, M. and Paprzycki, M., 2005. Ontology-based stereotyping in a travel support system', Proceedings of the XXI Autumn Meeting of Polish Information Processing Society, Wisla, Poland, pp.73–85.

[Nebel03]   Nebel, I-T., Smith, B. and Paschke, R., 2003. A user profiling component with the aid of user ontologies', Ws. Learning – Teaching – Knowledge – Adaptivity, Karlsruhe, Germany.

[Liu11]   Bing Liu, Web Data Mining Exploring Hyperlinks, Contents, and Usage Data. Springer-Verlag. 2011

[Chi02]   Ed H. Chi, Adam Rosien, and Jeffrey Heer. Lumberjack: Intelligent discovery and analysis of web user traffic composition. In Proc. of ACMSIGKDD Workshop on Web Mining for Usage Patterns and User Profiles. ACM Press, 2002.

[Strehl00]   Alexander Strehl, Joydeep Ghosh, and Raymond Mooney. Impact of similarity measures on web-page clustering. In Proceedings of the 17th National Conference on Artificial Intelligence: Workshop of Artificial Intelligence for Web Search (AAAI 2000), 30-31 July 2000, Austin, Texas, USA, pages 58–64. AAAI, July 2000.

[Agrawal93]   Rakesh Agrawal, Tomasz Imieli\&\#324;ski, and Arun Swami. 1993. Mining association rules between sets of items in large databases. In Proceedings of the 1993 ACM SIGMOD international conference on Management of data (SIGMOD '93), Peter Buneman and Sushil Jajodia (Eds.). ACM, New York, NY, USA, 207-216. DOI=10.1145/170035.170072

[Hahsler05]   Michael Hahsler and Bettina Grun and Kurt Hornik, arules - A Computational Environment for Mining Association Rules and Frequent Item Sets, Journal of Statistical Software. 2005

[Kliegr10]   Tomás Kliegr, Jan Rauch: An XML Format for Association Rule Models Based on the GUHA Method. RuleML 2010: 273-288

[Kliegr11]   Kliegr T., Svátek V, Ralbovský M., Šimůnek M. 2010. SEWEBAR-CMS: semantic analytical report authoring for data mining results. Journal of Intelligent Information Systems. 2011

[Dolejší04]   V. Dolejší, Jan Rauch , Milan Simunek, Vaclav Lin, The KL-Miner procedure for datamining, Neural Network World, 2004

[Rauch05]   Jan Rauch and Milan Simunek, GUHA method and granular computing, GrC 2005, IEEE

[Rauch09]   Jan Rauch and Milan Simunek, Action Rules and the GUHA Method: Preliminary Considerations and Results, ISMIS '09: Proceedings of the 18th International Symposium on Foundations of Intelligent Systems, 2009, Springer-Verlag

[Hashler11]   Hashler, 2011   http://michael.hahsler.net/research/bib/association_rules/, update: Wed Sep 14 14:50:36 2011

[Hipp00]   Jochen Hipp, Ulrich Guntzer, and Gholamreza Nakhaeizadeh. Algorithms for assocciation rule mining — a general survey and comparison. SIGKDD Explor. Newsl., 2(1):58–64, 2000.

[Zhang09]   Yuejin Zhang, Lingling Zhang, Guangli Nie, and Yong Shi. A survey of interestingness measures for association rules. Business Intelligence and Financial Engineering, International Conference on, 0:460–463, 2009.

[Nanavati01]   Amit A. Nanavati, Krishna P. Chitrapura, Sachindra Joshi, and Raghu Krishnapuram. Mining generalised disjunctive association rules. In CIKM '01, pages 482–489, New York, NY, USA, 2001. ACM.

[Antonie04]   Maria-Luiza Antonie and Osmar R. Zaıane. Mining positive and negative association rules: an approach for confined rules. In PKDD '04, pages 27–38, New York, NY, USA, 2004. Springer-Verlag New York, Inc.

[Srivastava00]   Srivastava, J., Cooley, R., Deshpande, M., & Tan, P.-N. (2000). Web usage mining: Discovery and applications of usage patterns from web data. SIGKDD Explorations , 1 (2), (pp 12-23).

[Bae03]   Bae, S. M., Ha, S. H., and Park, S. C., 2003. Fuzzy Web Ad Selector Based on Web Usage Mining. In IEEE Intelligent Systems 18, 6 , (pp 62-69).

[Kobsa01]   Kobsa, A., 2001. Tailoring privacy to users' needs. In UM '01: Proceedings of the 8th International Conference on User Modeling 2001 , (pp. 303-313). London, UK: Springer-Verlag.

[Mulvenna00]   Mulvenna, M. D., Anand, S. S., and Büchner, A. G., 2000. Personalisation on the Net using Web mining: introduction. Commun. ACM 43, 8 (pp 122-125).

[Zadeh08]   Zadeh, P. M. & Moshkenani, M. S. (2008). Mining Social Network for Semantic Advertisement. In Proceedings of the 2008 Third international Conference on Convergence and Hybrid information Technology - Volume 01 (November 11 - 13, 2008). ICCIT. IEEE Computer Society, Washington, DC, (pp 611-618).

[Tsatsou09]   D. Tsatsou, F. Menemenis, I. Kompatsiaris, P. C. Davis. A semantic framework for personalised ad recommendation based on advanced textual analysis. ACM International conference in recommender systems (RecSys) 2009, pp. 217-220

[Tsatsou11]   D. Tsatsou, S. Papadopoulos, I. Kompatsiaris, P.C. Davis, "Distributed Technologies for Personalised Advertisement Delivery", in Online Multimedia Advertising: Techniques and Technologies, Xian-Sheng Hua, Tao Mei, Alan Hanjalic (editors), IGI Global, 2011

[Burke07]   Burke, R., 2007. Hybrid web recommender systems. In The Adaptive Web, Lecture Notes in Computer Science, chap. 12, (pp. 377-408).

[Tziviskou07]   Christina Tziviskou and Marco Brambilla. 2007. Semantic personalisation of web portal contents. In Proceedings of the 16th international conference on World Wide Web (WWW '07). ACM, New York, NY, USA, 1245-1246. DOI=10.1145/1242572.1242788

[Amatriain09]   Xavier Amatriain, Josep M. Pujol, and Nuria Oliver. 2009. I Like It... I Like It Not: Evaluating User Ratings Noise in Recommender Systems. In Proceedings of the 17th International Conference on User Modeling, Adaptation, and Personalisation: formerly UM and AH (UMAP '09), Geert-Jan Houben, Gord Mccalla, Fabio Pianesi, and Massimo Zancanaro (Eds.). Springer-Verlag, Berlin, Heidelberg, 247-258. DOI=10.1007/978-3-642-02247-0_24

[Pannu11]   Pannu, M.; Anane, R.; Odetayo, M.; James, A.; , "Explicit user profiles in web search personalisation," Computer Supported Cooperative Work in Design (CSCWD), 2011 15th International Conference on , vol., no., pp.416-421, 8-10 June 2011. doi: 10.1109/CSCWD.2011.5960107

[Pazzani07]   Pazzani, M. J., & Billsus, D. (2007). D.: Content-based recommendation systems. In The Adaptive Web: Methods and Strategies of Web Personalisation. Volume 4321 of Lecture Notes in Computer Science , vol. 4321. (pp 325-341).

[Dumais03]   Susan Dumais, Edward Cutrell, JJ Cadiz, Gavin Jancke, Raman Sarin, and Daniel C. Robbins. 2003. Stuff I've seen: a system for personal information retrieval and re-use. In Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR '03). ACM, New York, NY, USA, 72-79. DOI=10.1145/860435.860451

[Lops11]   Pasquale Lops, Marco de Gemmis and Giovanni Semeraro, 2011. Content-based Recommender Systems: State of the Art and Trends. In Recommender Systems Handbook, Springer US, Part 1, 73-105, DOI: 10.1007/978-0-387-85820-3_3

[Ribeiro05]   Berthier Ribeiro-Neto, Marco Cristo, Paulo B. Golgher, and Edleno Silva de Moura. 2005. Impedance coupling in content-targeted advertising. In Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR '05). ACM, New York, NY, USA, 496-503. DOI=10.1145/1076034.1076119

[Chu09]   Chu, W., & Park, S. T., 2009. Personalised recommendation on dynamic content using predictive bilinear models. In WWW '09: Proceedings of the 18th international conference on World wide web, (pp. 691-700). New York, NY, USA: ACM.

[Papadopoulos09] Papadopoulos S., Menemenis F., Kompatsiaris Y., Bratu B. 2009. Lexical Graphs for Improved Contextual Ad Recommendation. In Proceedings of the 31st European Conference on Information Retrieval (Toulouse, France, April 2009).

[Zeng10]   Y. Zeng, Y. Wang, Z. Huang, D. Damljanovic, N. Zhong, and C. Wang. User interests: Definition, vocabulary, and utilization in unifying search and reasoning. In AMT, pages 98–107, 2010.

[VICON]   Vicon motion capture systems: http://www.vicon.com/products/viconmx.html

[SOFT]   SoftKinetic company: http://www.softkinetic.com/

[INTEL12]   Intel and SoftKinetic develop interfaces for interactive adds: http://techcrunch.com/2012/01/30/softkinetic-and-intel-partner-for-minority-report-style-ads/

[SAM12]   Samsung integrates cameras directly in their TVs: http://www.sananews.net/english/2012/01/samsung-will-include-integrated-camera-and-microphone-in-smart-tv-2-0/

[MIC_AVA12]   From real humans to avatars in real-time. http://www.microsoft.com/presspass/features/2012/jan12/01-03Future.mspx

[MIC_AMB11]   Ambient intelligence at Microsoft: http://www.engadget.com/2011/05/03/microsofts-home-of-the-future-lulls-teens-to-sleep-with-tweets/

[Ballendat10]   Ballendat, T., Marquardt, N., and Greenberg, S. Proxemic Interaction: Designing for a Proximity and Orientation- Aware Environment. Proc. of ITS'10, ACM (2010).

[Streitz03]   Streitz, N., et al. Ambient displays and mobile devices for the creation of social architectural spaces. In Public and Situated Displays. Kluwer, 2003, 387-410.

[Vogel04]   Vogel, D. et al. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. Proc. of UIST'04, ACM (2004).

[Greenberg11]   Greenberg, S., Marquardt, N., et al. Proxemic interactions:the new ubicomp? interactions 18, ACM (2011), 42–50.

[Hall66]   Hall, E.T. The Hidden Dimension. Doubleday, N.Y, 1966.

[Yang11]   Shuang-Hong Yang, Bo Long, Alex Smola, Narayanan Sadagopan, Zhaohui Zheng, and Hongyuan Zha. 2011. Like like alike: joint friendship and interest propagation in social networks. In Proceedings of the 20th international conference on World wide web (WWW '11). ACM, New York, NY, USA, 537-546. DOI=10.1145/1963405.1963481

[Zhang11]   Zhang, Yunlong Zhou, Jingyu Cheng, Jia. 2011. Preference-Based Top-K Influential Nodes Mining in Social Networks. IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), pp. 1512 - 1518.

[Sieg07]   Sieg A., Mobasher B., Burke R., 2007. Learning Ontology-Based User Profiles: A Semantic Approach to Personalised Web Search. In IEEE Intelligent Informatics Bulletin, Vol. 8, No.1.

[Trajkova04]   Trajkova, J., & Gauch, S., 2004. Improving ontology-based user profiles. In Proceedings of RIAO , (pp. 380-389).

[Kearney05]   Kearney, P., Anand, S.S., Shapcott, M., 2005. Employing a domain ontology to gain insights into user behaviour. In: Proceedings of the 3rd Workshop on Intelligent Techniques for Web Personalisation, at IJCAI 2005, Edinburgh, Scotland.

[Middleton04]   Middleton, S. E., Shadbolt, N. R., & De Roure, D. C., 2004. Ontological user profiling in recommender systems. ACM Transactions on Information Systems, 22 (1), (pp 54-88).

[Zhang07]   Hui Zhang, Yu Song, and Han-Tao Song. 2007. Construction of Ontology-Based User Model for Web Personalisation. In Proceedings of the 11th international conference on User Modeling (UM '07), Cristina Conati, Kathleen Mccoy, and Georgios Paliouras (Eds.). Springer-Verlag, Berlin, Heidelberg, 67-76. DOI=10.1007/978-3-540-73078-1_10

[VanderWal04]   Vander Wal, T. 2004. You Down with Folksonomy? Retrieved February 20, 2012 from http://www.vanderwal.net/random/entrysel.php?blog=1529

[Gemmis08]   Marco de Gemmis, Pasquale Lops, Giovanni Semeraro, and Pierpaolo Basile. 2008. Integrating tags in a semantic content-based recommender.

In Proceedings of the 2008 ACM conference on Recommender systems (RecSys '08). ACM, New York, NY, USA, 163-170. DOI=10.1145/1454008.1454036

[Cantador08]   Cantador, I., Szomszor, M., Alani, H., Fernández, M., & Castells, P., 2008. Enriching ontological user profiles with tagging history for multi-domain recommendations. In Proceedings of the 1st International Workshop on Collective Semantics: Collective Intelligence and the Semantic Web (CISWeb 2008), (pp. 5-19).

[Miller95]   Miller, G. A., 1995. WordNet: A Lexical Database for English. In Communications of the Association for Computing Machinery, 38(11), pp. 39-41.

[Sieg10]   Ahu Sieg, Bamshad Mobasher, and Robin Burke. 2010. Improving the effectiveness of collaborative recommendation with ontology-based user profiles. In Proceedings of the 1st International Workshop on Information Heterogeneity and Fusion in Recommender Systems (HetRec '10). ACM, New York, NY, USA, 39-46. DOI=10.1145/1869446

[Benitez01]   Ana B. Benitez , Di Zhong , Shih-Fu Chang , John R. Smith, MPEG-7 MDS Content Description Tools and Applications, Proceedings of the 9th International Conference on Computer Analysis of Images and Patterns, p.41-52, September 05-07, 2001

[Belk10]   Belk, M., Germanakos, P., Tsianos, N., Lekkas, Z., Mourlas, C., & Samaras, G., 2010. Adapting Generic Web Structures with Semantic Web Technologies: A Cognitive Approach. 4th International Workshop on Personalised Access, Profile Management, and Context Awareness in Databases (PersDB 2010)

[Bhowmick10]   P.K. Bhowmick, S. Sarkar and A. Basu, 2010. Ontology based user modelling for Personalised Information Access. International Journal of Computer Science and Applications, vol. 7, no. 1, pages 1–22.

[Weissenberg04] Norbert Weissenberg, Agnes Voisard, and Ruediger Gartmann. 2004. Using ontologies in personalised mobile applications. In Proceedings of the 12th annual ACM international workshop on Geographic information systems (GIS '04). ACM, New York, NY, USA, 2-11. DOI=10.1145/1032222.1032225

[Auer07]   Soeren Auer, Christian Bizer, Georgi Kobilarov, Jens Lehmann, Richard Cyganiak, and Zachary Ives. 2007. DBPedia: a nucleus for a web of open data. In Proceedings of the 6th international The semantic web and 2nd Asian conference on Asian semantic web conference (ISWC'07/ASWC'07), Karl Aberer, Philippe Cudre;-Mauroux, Key-Sun Choi, Natasha Noy, Dean Allemang, Kyung-Il Lee, Lyndon Nixon, Jennifer

Golbeck, Peter Mika, Diana Maynard, Riichiro Mizoguchi, and Guus Schreiber (Eds.). Springer-Verlag, Berlin, Heidelberg, 722-735.

[Gauch07]   S. Gauch, M. Speretta, A. Chandramouli, and A. Micarelli. User profiles for personalised information access. In P. Brusilovsky, A. Kobsa, and W. Nejdl, editors, The Adaptive Web, volume 4321 of Lecture Notes in Computer Science, pages 54–89. Springer Berlin / Heidelberg, 2007.

[Aquin10]   M. d'Aquin, S. Elahi, and E. Motta. Semantic monitoring of personal web activity to support the management of trust and privacy. In SPOT 2010: 2nd Workshop on Trust and Privacy on the Social and Semantic Web, 2010. Co-located with ESWC2010, the 7th European Extended Semantic Web Conference held 30 May-3 June 2010 at Heraklion (Greece).

[Fujimoto11]   H. Fujimoto, M. Etoh, A. Kinno, and Y. Akinaga. Web user profiling on proxy logs and its evaluation in personalisation. In X. Du, W. Fan, J. Wang, Z. Peng, and M. Sharaf, editors, Web Technologies and Applications, volume 6612 of Lecture Notes in Computer Science, pages 107–118. Springer Berlin / Heidelberg, 2011.

[Xu11]   G. Xu, Y. Zhang, L. Li, G. Xu, Y. Zhang, and L. Li. Web mining and recommendation systems. In Y. Zhang, editor, Web Mining and Social Networking, volume 6 of Web Information Systems Engineering and Internet Technologies, pages 169–188. Springer US, 2011.

[Kelly03]   Kelly, D., Teevan, J.: Implicit feedback for inferring user preference: a bibliography. ACM SIGIR Forum 37(2) (2003) 18-28

[Zhang10]   Bangzuo Zhang; Yu Guan; Haichao Sun; Qingchao Liu; Jun Kong; , "Survey of user behaviors as implicit feedback," Computer, Mechatronics, Control and Electronic Engineering (CMCE), 2010 International Conference on , vol.6, no., pp.345-348, 24-26 Aug. 2010

[Oard98]   D. W. Oard, J. Kim, 1998. Implicit Feedback for Recommender Systems. In AAAI Workshop on Recommender Systems, Madison, WI, pp. 81-83.

[Fujimoto11b]   Hiroshi Fujimoto, Minoru Etoh, Akira Kinno and Yoshikazu Akinaga, Topic Analysis of Web User Behavior Using LDA Model on Proxy Logs, ADVANCES IN KNOWLEDGE DISCOVERY AND DATA MINING, 2011

[Holub10]   Michal Holub and Maria Bielikova. 2010. Estimation of user interest in visited web page. In Proceedings of the 19th international conference on World wide web (WWW '10). ACM, New York, NY, USA, 1111-1112.

[Ohmura11]   Hayato Ohmura, Teruaki Kitasuka, and Masayoshi Aritsugi. 2011. Web browsing behavior recording system. In Proceedings of the 15th international conference on Knowledge-based and intelligent information and engineering systems - Volume Part IV (KES'11), Andreas König,

Andreas Dengel, Knut Hinkelmann, Koichi Kise, and Robert J. Howlett (Eds.), Vol. Part IV. Springer-Verlag, Berlin, Heidelberg, 53-62.

[Jakobsson08]   Markus Jakobsson, Ari Juels, and Jacob Ratkiewicz. Privacy-Preserving history mining for web browsers. In Web 2.0 Security and Privacy, Oakland, CA, May 2008. IEEE, IEEE.

[Ghorab09]   M. Rami Ghorab, Johannes Leveling, Dong Zhou, Gareth J. F. Jones, and Vincent Wade. 2009. Identifying common user behaviour in multilingual search logs. In Proceedings of the 10th cross-language evaluation forum conference on Multilingual information access evaluation: text retrieval experiments (CLEF'09), Carol Peters, Giorgio Maria Di Nunzio, Mikko Kurimo, Thomas Mandl, and Djamel Mostefa (Eds.). Springer-Verlag, Berlin, Heidelberg, 518-525.

[Park11]   Kinam Park, Hyesung Jee, Taemin Lee, Soonyoung Jung and Heuiseok Lim, 2011. Automatic extraction of user's search intention from web search logs. Multimedia Tools and Applications (2011)

[Facca05]   Federico Michele Facca and Pier Luca Lanzi. Mining interesting knowledge from weblogs: a survey. Data & Knowledge Engineering, 53(3):225–241, 2005. ISSN 0169-023X. doi: http://dx.doi.org/10.1016/j.datak.2004.08.001

[Granka04]   Laura A. Granka, Thorsten Joachims, and Geri Gay. Eye-tracking analysis of user behavior in www search. In SIGIR '04: Proceedings of the 27th annual interna- tional ACM SIGIR conference on Research and development in information retrieval, pages 478–479, New York, NY, USA, 2004. ACM Press. ISBN 1-58113-881-4. doi: http://doi.acm.org/10.1145/1008992.1009079.

[Barla06]   Michal Barla, 2006. Interception of user's interests on the web. In Vincent P. Wade, Helen Ashman, and Barry Smyth, editors, AH, volume 4018 of Lecture Notes in Computer Science, pages 435–439. Springer, 2006. ISBN 3-540-34696-1.

[Thomas03]   Richard Thomas, Gregor Kennedy, Steve Draper, Rebecca Mancy, Murray Crease, Huw Evans, and Phil Gray. Generic usage monitoring of programming students, 2003.

[Kim06]   Il Kim, Bong Joon Choi, and Kyoo Seok Park. A study on web-usage mining control system of using page scroll. In Sven F. Crone, Stefan Lessmann, and Robert Stahlbock, editors, DMIN, pages 337–342. CSREA Press, 2006. ISBN 1-60132-004-3.

[Horrocks04]   I. Horrocks and P.F. Patel-Schneider. Reducing OWL entailment to Description Logic satisfiability. J. Web Sem., 1(4):345{357, 2004.

[Dalakleidi11] K. Dalakleidi, S. Dasiopoulou, G. Stoilos, V. Tzouvaras, G. Stamou, I. Kompatsiaris, 2011. Semantic Representation of Multimedia Content, in Knowledge-Driven Multimedia Information Extraction and Ontology Evolution, G. Paliouras, C. Spyropoulos and G. Tsatsaronis (Eds), Springer Verlang.

[Li02] Yuefeng Li and Y. Y. Yao. 2002. User Profile Model: A View from Artificial Intelligence. In Proceedings of the Third International Conference on Rough Sets and Current Trends in Computing (TSCTC '02), James J. Alpigini, James F. Peters, Jacek Skowronek, and Ning Zhong (Eds.). Springer-Verlag, London, UK, UK, 493-496.

[Völker11] Johanna Völker and Mathias Niepert. 2011. Statistical schema induction. In Proceedings of the 8th extended semantic web conference on The semantic web: research and applications - Volume Part I (ESWC'11), Grigoris Antoniou, Marko Grobelnik, Elena Simperl, Bijan Parsia, and Dimitris Plexousakis (Eds.), Vol. Part I. Springer-Verlag, Berlin, Heidelberg, 124-138.

[Rizzo11] G. Rizzo and R. Troncy, 2011. Nerd: evaluating named entity recognition tools in the web of data. In Workshop on Web Scale Knowledge Extraction (WEKEX'11), 2011. 21

[Sheth01] Sheth, M., Avant, D., Bertram, C., 2001. System and method for creating a semantic web and its applications in browsing, searching, profiling, personalisation and advertising, US Patent Application 6311194, October 2001.

[Ardissono03] Ardissono, L., Gena, C., Torasso, P., Bellifemine, F.,Chiarotto, A., Difino, A. and Negro, B., 2003. Personalisedrecommendation of TV programs, Proceedings of the Int.Conf. on Advances in Artificial Intelligence, Pisa, Italy, pp.474–486.

[Ardissono04] Ardissono, L., Gena, C., Torasso, P., Bellifemine, F., Difino, A.and Negro, B., 2004. User Modelling and Recommendation Techniques for Personalised Electronic Program Guides, Personalised Digital Television: Targeting Programs toIndividual Viewers, Springer, The Netherlands.

[Adomavicius08] Gediminas Adomavicius and Alexander Tuzhilin. 2008. Context-aware recommender systems. In Proceedings of the 2008 ACM conference on Recommender systems (RecSys '08). ACM, New York, NY, USA, 335-336. DOI=10.1145/1454008.1454068

[Palmisano08] Cosimo Palmisano, Alexander Tuzhilin, Michele Gorgoglione, 2008. Using Context to Improve Predictive Modeling of Customers in Personalisation Applications. IEEE Trans. Knowl. Data Eng. 20(11): 1535-1549.

[Martinez09] Ana Belen Barragans Martinez, Jose J. Pazos Arias, Ana Fernandez Vilas, Jorge Garcia Duque, Martin Lopez Nores, 2009. What's on tv

tonight? An efficient and effective personalised recommender system of TV programs. icce, pp.1-2, 2009 Digest of Technical Papers International Conference on Consumer Electronics, 2009

[Siskos05]   Siskos, Y., 2005. UTA Methods. Multiple Criteria Decision Analysis: State of the Art suveys, Springer, New York 2005, 297-334.

[Dembczyński10] Krzysztof Dembczyński, Wojciech Kotłowski, Roman Słowiński and Marcin Szelag, 2010. Learning of Rule Ensembles for Multiple Attribute Ranking Problems. In Preference Learning (Johannes Fürnkranz and Eyke Hüllermeier, editors), pages 217-247, Springer-Verlag.

[Greco08]   Salvatore Greco, Vincent Mousseau, Roman Slowinski, 2008. Ordinal regression revisited: Multiple criteria ranking using a set of additive value functions. European Journal of Operational Research 191(2): 416-436.

[Grycza04]   Grycza, 2004. A generalized regression approach to multiple-criteria decision support with an additive utility function. Diploma thesis, supervisor : prof. Slowinski, Poznan University of Technology, 2004.

[Abrams07]   Abrams Z. and Vee E., 2007. Personalised Ad Delivery when Ads Fatigue: An Approximation Algorithm. In Proc. Workshop on Internet and Network Economics (WINE) (pp. 535-540).

[Resnick94]   Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., Riedl, J.: Grouplens: An open architecture for collaborative filtering of netnews. In: Conference on Computer Supported Cooperative Work, pp. 175–186. ACM Press, New York (1994)

[Breese98]   Breese, J., Heckerman, D., Kadie, C.: Empirical analysis of predictive algorithms for collaborative filtering. In: 14th Conference on Uncertainty in Artificial Intelligence, pp. 43–52. Morgan Kaufman, San Francisco (1998)

[Resnick94]   Pennock, D., Horvitz, E., Lawrence, S., Giles, C.L.: Collaborative filtering by personality diagnosis: A hybrid memory- and model-based approach. In: 16th Conference on Uncertainty in Artificial Intelligence, pp. 473–480 (2000)

[Kleinberg04]   Kleinberg, J., Sandler, M.: Using mixture models for collaborative filtering. In: 36th ACM Symposium on Theory Of Computing, pp. 569–578. ACM Press, New York (2004)

[Kelleher03]   Kelleher, J., Bridge, D.: Rectree centroid: An accurate, scalable collaborative recommender. In: Cunningham, P., Fernando, T., Vogel, C. (eds.) 14th Irish Conference on Artificial Intelligence and Cognitive Science, pp. 89–94 (2003)

[Gruber05]   Gruber, T.: Ontology of folksonomy: A mash-up of apples and oranges. MTSR2005 (2005)

[Mika07]   Mika, P.: Social Networks and the Semantic Web. Springer (2007)

[Kim07]   Kim, H., Yang, S., Song, S., Breslin, J. G., Kim, H.: Tag Mediated Society
                 with SCOT Ontology. ISWC2007. (2007)

[Passant08]   Passant, A., Laublet, P.: Meaning Of A Tag: A Collaborative Approach to
                 Bridge the Gap Between Tagging and Linked Data. LDOW2008 (2008)

[Kliegr09]   Kliegr, T. , 2009. UTA - NM: Explaining stated preferences with additive
                 non-monotonic utility functions (0). In Preference Learning (PL-09)
                 ECML/PKDD-09 Workshop.

[Svátek05]   Vojtech Svátek, Jan Rauch, Martin Ralbovský: Ontology-Enhanced
                 Association Mining.EWMF/KDO 2005: 163-179

[Marinica10]   C. Marinica et F. Guillet, 2010. *Filtering Discovered Association Rules
                 Using Ontologies*. IEEE Transactions on Knowledge and Data Engineering
                 (TKDE) Journal, special issue Domain-driven Data Mining, vol. 22, no. 6,
                 pp. 784-797.

[Jacquet82]   Jacquet-Lagrèze, E. and Y. Siskos (1982). Assessing a set of additive
                 utility functions for multicriteria decision making: The UTA method,
                 European Journal of Operational Research, 10 (2), 151–164.

[Bobillo08]   Bobillo, F.; Straccia, U.; , "fuzzyDL: An expressive fuzzy description logic
                 reasoner," Fuzzy Systems, 2008. FUZZ-IEEE 2008. (IEEE World
                 Congress on Computational Intelligence). IEEE International Conference
                 on , vol., no., pp.923-930, 1-6 June 2008

[Kobsa85]   Kobsa, A. (1985). Benutzermodellierung in Dialogsystemen (User
                 Modeling in Dialog Systems), Volume 115 computer science lecture note.
                 Springer,   Berlin,   Heidelberg.   http://www.ics.uci.edu/~kobsa/kobsa-
                 publi.htm

[Koren08]   Yehuda Koren. 2008. Factorization meets the neighborhood: a
                 multifaceted collaborative filtering model. In Proceedings of the 14th ACM
                 SIGKDD international conference on Knowledge discovery and data
                 mining (KDD '08). ACM, New York, NY, USA, 426-434.

[Balasundaram11] Balabhaskar Balasundaram, Sergiy Butenko, and Illya V. Hicks. 2011.
                 Clique Relaxations in Social Network Analysis: The Maximum k-Plex
                 Problem. Oper. Res. 59, 1 (January 2011)

[Taylor02]   A. Taylor, R. Harper Switching on to switch off: a analysis of routine TV
                 watching habits and their implications for electronic programme guide
                 design, 2002

[Isobe03]   T. Isobe, M. Fujiwara, H. Kaneta, N. Uratani, T. Morita, Development and
                 Features of a TV Navigation System, IEEE 2003

[Zhang05]   H. Zhang, S. Zheng, Personalized TV Program Recommendation based
                 on TV-Anytime Metadata, IEEE, 2005

[Huet05]   B. Huet, J. Jiten, B. Merialdo, Personalization of hyperlinked video in interactive television, IEEE, 2005

[Shin09]   H. Shin, M. Lee, E. Yi Kim, Personalized Digital TV Content Recommendation with Integration of User Behaviour Profiling and Multimodal Content Rating, IEEE Transactions on Consumer Electronics, 2009

[Song12]   S. Song, H. Moustafa, H. Afifi, A Survey on Personalized TV and NGN Services through Context-Awareness, 2012

[Dimitrova03]   N. Dimitrova, J. Zimmerman, A. Janevski, L. Agnihotri, N. Haas, R. Bolle, Content Augmentation Aspects of Personalized Entertainment Experience, In 3rd Workshop on Personalization in Future TV, 2003

[Yu06]   Z. Yu, X. Zhou, D. Zhang, C. Chin, X. Wang, J. Men, Supporting Context-Aware Media Recommendations for Smart Phones, IEEE 2006

[Harrison08]   Chris Harrison, Brian Amento, and Larry Stead. 2008. iEPG: an ego-centric electronic program guide and recommendation interface. In Proceedings of the 1st international conference on Designing interactive user experiences for TV and video (UXTV '08). ACM, New York, NY, USA, 23-26.

[Aguzzoli02]   Stefano Aguzzoli, Paolo Avesani, and Paolo Massa. 2002. Collaborative Case-Based Recommender Systems. In Proceedings of the 6th European Conference on Advances in Case-Based Reasoning (ECCBR '02), Susan Craw and Alun D. Preece (Eds.). Springer-Verlag, London, UK, UK, 460-474.

[CSME11]   Skyware Smart Media Engine, Karsten Trint, release: 10.2011, http://www.condat.de/portfolio/medien/broadcast/zweiseiter_skyware_smart_media_engine.pdf

[Ju08]   Ju, W., Lee, B.A., and Klemmer, S.R. Range: exploring implicit interaction through electronic whiteboard design. Proc. of CSCW '08, ACM (2008), 17-26.

[Vinciarelli09] A.Vinciarelli, M.Pantic and H.Bourlard Social Signal Processing: Survey of an Emerging Domain Image and Vision Computing Journal, Vol. 27, No. 12, pp. 1743-1759 (2009).