Contextualised user profiling in networked media environments

Dorothea Tsatsou[1], Lyndon Nixon[2], Matei Mancas[3], Miroslav Vacura[4], Rüdiger Klein[5], Julien Leroy[3], Jaroslav Kuchař[4], Tomáš Kliegr[4], Manuel Kober[5], Maria Loli[1], Vasileios Mezaris[1]

[1] CERTH-ITI, Thessaloniki, Greece
{dorothea,mloli,bmezaris@iti.gr}
[2] STI International, Vienna, Austria
lyndon.nixon@sti2.org
[3] University of Mons, Mons, Belgium
{matei.mancas,julien.leroy@umons.ac.be}
[4] University of Economics Prague, Prague, Czech Republic
{vacuram,jaroslav.kuchar, tomas.kliegr@vse.cz}
[5] Fraunhofer IAIS, Sankt Augustin, Bonn, Germany
{ruediger.klein,manuel.kober@iais.fraunhofer.de}

Abstract. TV and Web convergence is becoming more and more a reality. This paper provides an overview of the opportunities and challenges that arise in future TV environments regarding unobtrusive, context-aware personalisation of digital media content. Subsequently, it describes the vision and first conceptual personalisation approach within the LinkedTV EU project. LinkedTV aims to seamlessly interlink TV and Web content, while enhancing the users' TV experience by producing contextualized user models based on the online transactional and physical behaviour of the user.

Keywords: semantic user profiling, contextualisation, behavioural tracking, digital media, future TV.

1 Introduction

A future media landscape is envisioned in the Web of Networked Media, in which online devices are ubiquitous and digital media is interwoven, in order to facilitate access and retrieval over media content. In this environment, users will interact with any media at any time with any device, effortlessly switching between content, thus changing the landscape of future TV and enabling new opportunities in targeted media content delivery. Two pivotal factors supporting this vision are the emergence of Social TV and the technological advances enhancing user-TV interaction. Social TV allows users to share TV views/likes/recommendations on the Web and combine TV viewing with Social Media communication on or alongside the TV, while cameras and microphones embedded in future TVs support gesture and voice control, thus permitting the incorporation of viewer behaviour in personalised TV environments.

The convergence of TV and the Web has opened a whole new perspective in personalised systems by offering the possibility to interlink behaviour with respect to multidisciplinary content [1] for consumer usage, whether that might be audio, video or text (transcripts, articles, tags, comments etc), while tailoring it to the user's needs and concrete situation. Thus a 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 also render them scalable and usable for subsequent matching to the disparate media content.

The **Television Linked To The Web (LinkedTV)** [1] project aims to provide a novel practical approach to future networked media, where audiovisual content is decoupled from place, device or source. The union of television with the Web means that television content and Web content should and can be seamlessly connected, while browsing TV and Web content should be interrelated in such a way that eventually even "surfing the Web" or "watching TV" will become a meaningless distinction. To this end, people will be allowed to seamlessly delve into and browse content within the programme itself at the level of individual objects on screen or things which are mentioned or referred to. Within the scope of LinkedTV, triggerable video objects will be connected to additional content and information, thus augmenting the features of the audiovisual content with more comprehensive descriptions.
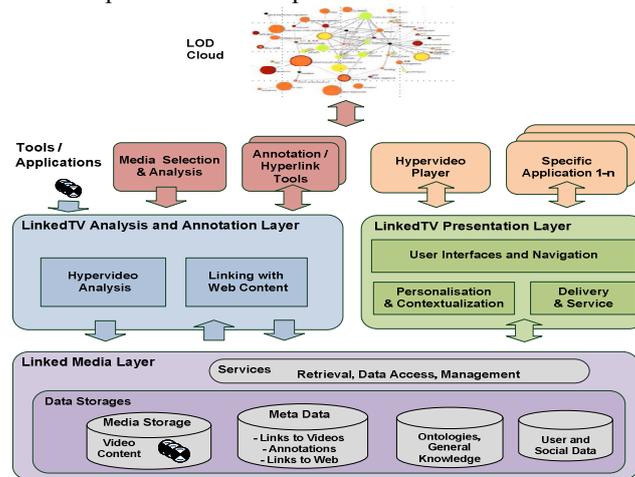

Fig. 1. The LinkedTV architecture overview

Nevertheless the general issue pertaining to advanced personalisation in networked media platforms, such as LinkedTV, is capitalizing from the information available in user-consumed digital data by understanding and unifying information from various heterogeneous sources such as video, audio, text and social media, based on the user's transactional and reactional behaviour with such content. In addition, LinkedTV aims to extract different user contexts based on two perspectives: a) on-platform interaction with the content, depending on concrete conditions such as time, location and actions and b) behavioral tracking in order to determine the physical state (alone, with company) and reaction to the content (mood, attention).

The rest of the paper is structured as follows: Section 2 presents an overview of transaction-based user modeling strategies for digital media environments. Section 3 outlines methodologies for tracking the physical behaviour of a user in camera-enhanced environments. Section 4 explores the benefits of semantic profiling for aligning heterogeneous user information resources. Section 5 will present the initial conceptual personalisation and contextualisation approach in LinkedTV. Finally Section 6 concludes with the outlook for the realization of an advanced networked media platform such as LinkedTV.

---

[1] www.linkedtv.eu; FP7 IST IP, 2011-2015. Twitter @linkedtv

2 User Modeling based on user transactions

In a personalised environment, information about the user can be either explicitly defined by the user or implicitly extracted through his interaction with the system and relevant resources. Earlier studies have shown that in general users are reluctant to disclose sensitive personal information on the web [2], especially in terms of their demographic data, or even indisposed to put in the effort to provide preference-specific information. However, most social networks and other online services nowadays allow and prompt users to provide initial explicit information about themselves, such as biographic data, their preferences, and their links to other people, thus providing a first insight over a user.

This explicit information can be used to form categories *"constituting strong points of commonalities"* [3] among users called stereotypes. Stereotypes can serve as a primitive clustering methodology that can be used to expand the initial knowledge about the user according to information available about "similar" users.

Nonetheless, due to the obtrusive 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. Implicitly capturing and managing user information gives rise to different privacy compromising issues that an intelligent personalisation system must address. In addition, some online sources are not reliable. That is why some partners of LinkedTV decided to use white-listing of online sources to ensure required level of reliability of additional information obtained from the web.

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. Data can be collected on the client machine (client-side), by the application server itself (server-side), or both. Client side monitoring is deemed more convenient and can provide greater granularity about a user's interaction with a web platform. Server side monitoring has several important advantages, such as unobtrusiveness, invisibility, speed and possibility for interlinking content, but never gained much popularity since it suffers from accuracy problems. Foremost, handling personal data on the server side seriously compromises user privacy.

The sources of implicit transactional information pertaining to a user are three-faceted, involving learning predictive user models from a) data extracted from a user's the content consumption history (i.e. user-assigned ratings [4], data/text mining [5]), b) identifying similarities with peers in a social information exchange environment (i.e. collaborative filtering [6], social network mining [7]) and c) taking into account predefined knowledge encompassed in ontological structures [8].

These aspects come with advantages and challenges. The content-based approach is the most straightforward and indicative of a user's individual behaviour, while peer-based approaches may provide insight about foreseen preferences that a user has not yet expressed. Both these methods though are hampered by the cold-start problem [9] when new users are introduced on a platform, as well as by data sparcity and scalability problems [6]. Knowledge-based approaches resort on a finite (therefore arguably scalable), uniform vocabulary [10] that can augment user-induced information with world knowledge but are hindered by the lack of mappings between free-form information and the ontological knowledge. 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 [11]. To this end, LinkedTV aims to follow a hybrid approach where background semantic knowledge

will serve as the basis for understanding users based on their transactions with both the content as well as peers in social networks, as will be further analysed in section 4, while imparting initial information that can leverage the lack of data about new users. It will also provide the means to unify, structure and reduce the load of user-related information that builds up over time from the user's transactional data.

Meanwhile, the majority of existing contextualisation approaches rely on segmenting user behavior based on factors like time, location and company of other people [12]. Beyond this semi-explicit information, the most straightforward contextual influence obviously comes from recent content consumptions. 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 [13].

In digital media in particular, the metadata desultory and sparcity in heterogeneous multimedia content is the prevailing challenge in understanding and capturing user preferences and context. In effect, some of the current personalisation approaches in TV programme recommendation still avoid cumbersome digital data analysis and rely on user-defined preferences [1], while others [14] attempt to enrich TV program metadata based on the electronic program guide (EPG) data and concept dictionaries. In many media-oriented platforms, the current viewed context itself is the sole factor considered to provide additional content to a user [15]. Other TV programme recommendation approaches also consider the specific capabilities that a dedicated multimedia platform offers (actions over media playback). For instance [16] take into account the level of interest of the user, based both on usage history and the current actions (play, stop, rewind etc) performed on the player, and also the duration of the action and the consumed content metadata.

The contextualisation approach considered in LinkedTV will take into account temporal attributes (time of day, season etc), location information and the specific sub-topic that the user is currently or recently engaged in. This transactional information will be further enhanced by the physical features described in section 3.

The recently finished NoTube EU project[2] in particular also aimed at bringing Web and TV closer together via shared data models and content. NoTube focused on the recommendation of TV programs as atomic units, based on enrichment of the EPG data. LinkedTV goes further because it works on the level of fragments of both TV and Web content, and interlinks them. So the content needs analysis and annotation at both spatial and temporal level, and personalisation is now about personalising the choice of fragments of the TV content and the Web content to link to/from. Hence, only the content of most relevance to the user is linked to and from a clickable object on the video. In such an environment, apart from the focused contextual information available in the media fragments, the personalisation task may benefit from understanding additional user actions on the player such as skip or bookmark a media fragment.

Following the identification of user preferences, the goal of preference learning is to induce predictive preference models from empirical data [17]. 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 types of machine learning techniques oriented towards addressing this challenge, including (to list just a few that are mostly relevant to LinkedTV given the digital data available to the platform): learning preference relations (for instance via learning association rules between preferences), learning utility functions to rank object preferences and user clustering.

---

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

3   User Modeling based on user physical behaviour

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. A user's physical behaviour can be extracted in various ways from different kinds of input data. In this section, we will focus on audio-visual features and on behavioural tracking which falls under the scope of computer vision technologies. Moreover, we focus on applications concerning TV or home related experiences where the state of the art is much less deployed and on implicit behaviour analysis of the user physical reaction to the presented content, which means that explicit gestures for interfaces control are not in the scope of this section.

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 [18]. 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. Also other 3D camera-based explicit interfaces developed for TVs and interactive adds [19] 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 [20]. 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 sex or age can already be inferred from camera input. Some interesting implicit interfaces were developed using the notion of proxemics [21] in multi-user environments. The idea is that people's physical distance is correlated to social distance. The authors of [21] noticed several areas surrounding people that suggest certain types of interaction:

- Intimate distance: a really close space with high probability of physical contact.
- Personal distance: distance for interacting with relatives like family members or good friends.
- Social distance: 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: 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.

[22] identified five essential dimensions as a first-order approximation of key proxemics measures that should be considered: a) orientation (relative angles between entities), b) distance (between people, objects, and digital devices), c) motion (changes of distance and orientation over time), d) identity (of a person, or a particular device) and

e) location (as in the setup of environmental features such as walls and doors or movable furniture).

[23] 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. Fig. 2 which is extracted from [23] shows the different displays on the TV during the observer arrival. This system was initially designed to be used with precise motion capture systems [18], but part of it works also using low-cost depth cameras like the Kinect sensor.

Fig. 2. 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 [23].

As it can be seen here, the systems applying implicit interaction and behavioral tracking in TV setups are limited in number. As the market (both for depth cameras and TV manufacturers) is gradually oriented towards explicit and implicit interaction based on people's behavior, in LinkedTV we propose to develop a system which is fully adapted to different TV setups and which integrates with other profiling technologies.

To this end, within LinkedTV we plan to use a Kinect sensor which is low-cost and common to extract RGB, depth and audio features in order to determine user behaviour. 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. We will extend the [23] approach based on Greensberg's [22] five features:

- Location: determine the environment close to the TV. During time, the system will be able to learn where people come from and where 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. By observing people's activity over time, it is possible to find repetitive patterns and localize those patterns in space: in that way depending on people's position and features, it is possible to predict what they will do (based on what they are used to do in the same conditions).
- Identity and context: a person can be recognized by using several features; a combination of face recognition and biometric features (tall, small, fat, gender…) seems to be the better choice to capture identity and context. This point can also serve at discovering the number of people, if they are already known or not, and to extract biometric features mainly about their age and gender.
- Orientation: detect if the focus of attention is on the TV or not based on the direction of the body of the user. Audio features will also be used to see if there might be

a conversation between people, or if someone is talking. This feature is of high importance to refine the profile of the viewer.

- Distance and static features: 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's silhouettes like their symmetry (how symmetric are the hands, is there one hand pointing the TV screen …) and the contraction index (are the hands along the body or not).
- 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. Audio features can also provide clues 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. The motion features can also help in context change detection. For instance, a change in the user's behavior can happen due to the display of emotional content, but also to the arrival of friends, or other changes in viewer context.

Obviously, in the case of reactive behavioural tracking very distinct and sensitive user information is utilized. Apart from the required permission to capture and process the reactional behaviour of a user, a privacy preserving personalisation system should take care not to store rough (camera) data, as well as to ensure anonymisation of behavioural user features and secure server-client data communications. In addition, a social constraint arises with regard to third-party users (i.e. family members, guests) that are potentially unaware of the capabilities of such behavioural tracking mechanisms or reluctant to share physical personal information when in the same environment (e.g. a friend's house) where such a platform is active.

4 Semantic profiling for augmented user modeling

The information overload, instigated in future TV environments from the diversity and vastness of information in multimedia content, can be efficiently managed by the representation of the information in compact, lightweight ontological conceptualizations. Semantic profiling enables fine-grained structuring of user profiles on one hand and provides a way to represent uniform and meaningful conceptualizations of the diverse, ambiguous and sparsely described multimedia content on the other hand. By the term "semantic" in the context of LinkedTV project we refer to paradigms that employ methods and approaches based on formal ontologies as a way for building structured semantic user model descriptions.

To this end, within LinkedTV we will explore an approach to extract and express the semantics of user behaviour with respect to the viewed and consumed content, the user's reaction and physical disposal towards that content and his interactions with peers and express them in a uniform ontological vocabulary. 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.

Therefore several issues need to be dealt with in the context of implicitly capturing and representing a semantic user profile: a) determination of suitable ontology/ies that can provide a meaningful knowledge base potent to capture domain and user-pertinent semantics and of 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), b) the means to understand the semantics of user behaviour, i.e. map it to available knowledge and c) 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 and manageable by an intelligent inferencing engine.

There are several ontology languages and schemata (RDFS[3], OWL[4], SKOS[5], FOAF[6] etc) 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. Several general-purpose, cross-discipline ontologies (e.g. DBPedia [24]) can provide a core framework for understanding and representing broad, multidiscipline user preferences and augmenting them with aggregated information available on the Web such as Linked Open Data[7] over a wide range of domains. However, their limited expressivity de facto cannot support expression of more complex background information like domain-specific ontologies can. Furthermore, since the focus of LinkedTV is the domain of broadcasting, it is possible to employ domain-specific reference ontologies[8] with better structured semantics, available for re-use and alignment within multidisciplinary environments.

Semantic classification addresses the challenge of understanding the vastly heterogeneous multimedia data and interpreting the digital traces of a user, while expressing it in a holistic, machine-understandable format, i.e. mapping the content and its impact to the user to available ontological knowledge. Video content-related interpretation will be addressed within LinkedTV through intelligent audiovisual analysis (which is out of the scope of this paper) and enrichment of multimedia content metadata to support concept identification in hypervideo. Metadata enrichment can be achieved by employing existing natural language processing tools (e.g. OpenCalais[9], GATE[10], SProUT[11]) and extending their capabilities with new or integrating approaches as well as with Web mining over completely unstructured data.

Most notably, metadata analysis in LinkedTV will focus on the development of a *Linked Media Layer* – composed of a fine-grained mechanism for interlinking fragments of multimedia and annotation schemas. This mechanism is again based on a semantic approach for representation of media data. The Linked Media Layer will also integrate the media fragment specification and the metadata schemas together with the annotations gathered from web mining processes in order to provide a powerful search and retrieval module.

Another important classification (and unification) challenge regards interpretation of free-form content relevant to a user (textual descriptions, social activities, physical behaviour) that cannot be expected to be semantically pre-described (in contrast to annotated AV content), especially not in a common vocabulary. This involves looking into approaches for lexical analysis of textual content (e.g. [10]) to extract semantic mappings between raw text and ontological knowledge, as well as methodologies to track and understand a user's social web activities (e.g. NoTube's Beancounter tool). These methodologies will also support preference weighting, the weights denoting the level of interest or disinterest of extracted conceptual preferences based on such factors

---

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

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

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

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

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

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

[9] http://www.opencalais.com/

[10] http://gate.ac.uk/

[11] http://sprout.dfki.de/

such as the frequency and age of the preference. Physical behaviour interpretation remains mostly unexplored and in the scope of LinkedTV will initially reflect/adapt the impact of a content-extracted preference to the user (positive/negative, degree of impact). Later on, the semantic representation of distinct behavioural features will be considered, given the existence of corresponding semantic knowledge in the reference ontology.

Finally, 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. Between known modelling schemata (e.g. activation of interests over a full ontology [8], vector space preference aggregation [25]) a rule-based user model is opted for LinkedTV with respect to addressing three main issues: i) scalability, especially with respect to the diverse and broad preferences emerging in a multimodal linked media environment, ii) ability to express complex patterns in user behaviour, most prominently in different contextual situations where different relations might pertain to the same concepts and iii) ability to express disinterests.

5   The LinkedTV conceptual personalisation approach

Given the need to take advantage of all the various sources of information derived from heterogeneous resources and extended user activities (physical behaviour in addition to content consumption) in a networked media platform, semantic user profiling is deemed as the predominant strategy for personalizing user preferences in LinkedTV. To this end, semantic user profiles can sufficiently address the intricacies posed within networked media environments by reducing the dimensionality and scale of digital media data with an interest in keeping user models lightweight and efficiently meaningful (under a "say less, mean more" theme), while being able to exploit the capabilities of various machine learning mechanisms for complex preference learning as well as the capacities of inference engines.

Furthermore, with the coalition of behavioural tracking strategies, LinkedTV aims to go a step further in contextualizing user preferences 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.

The question is how 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. LinkedTV's contextualised user model will constitute of multiple contextualised sub-models pertaining to different user situations based on his long-term preferences and related context-defining features.
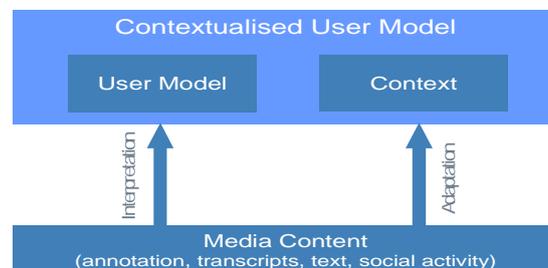
Fig. 3. High-level conceptualisation of LinkedTV's contextualised user model

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 concept & content recommendation, presentation of user adapted versions of a video and media search. While the contextualised profiling schema considered in LinkedTV will be implemented at a semantic level, thus rendering it adaptable to all three of the aforementioned use cases, since LinkedTV is focused towards linking media content with web content, the first use case is highlighted as the target focus for the CUM.

As aforementioned, in order to cope with the diversity and discrepancy of user-pertinent information in a multimodal networked media environment, such as LinkedTV, 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, personalisation and contextualisation strategies in LinkedTV aim to:

- Understand what the data pertinent to a user denotes, based on his transaction with and reaction to diverse content resources, including:
  - Unobtrusively extract information about the user, evaluate their nature (interest vs. disinterest) and estimate their impact to the user (weighted preferences)
  - Analyse that information and acquire meaningful semantic knowledge about the user.
  - Align disparate data stemming from content consumption, interactions in social networks, predefined domain knowledge, and explicitly defined user preferences.
- Augment user preferences with additional related semantic information based on:
  - World knowledge incorporated in ontological knowledge bases
  - Exploiting the rich information freely available in linked open data
  - Web mining in structured and unstructured resources
  - Exploiting semantic linkage of digital media in a Linked Media Layer
- Put acquired knowledge down in a uniform machine-understandable vocabulary, considering:
  - Using and maintaining expressive and lightweight ontological knowledge base/s
  - Producing and maintaining a compact and meaningful semantic (ontology-based) user model
- Learn user behaviour by:
  - Estimating preference ranking
  - Discovering relations between preferences (behavioural patterns)
- Harvest mass intelligence via:
  - Factorisation of peer-based user information stemming from social media interactions
  - Clustering users and learning group-specific knowledge
- Determine the behavioural (reactional) and physical situation of the user in order to provide 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).

- Understand behavioural information and adapt it the user profile by defining multiple contextualised facets of the user model.

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. 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. To this end, while initial approaches will be deployed 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 the processing techniques, thus osculating the possibility of minimizing server-client transmission of sensitive data. 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 the LinkedTV personalisation system competitive and suitable for commercial exploitation.

6   Conclusions

Current personalisation approaches in networked media environments are oriented towards unobtrusively extracting and understanding implicit user digital traces and translating them 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/her social networks activity and/or available semantic background knowledge. Moreover, information from embedded cameras can help trace the physical state of the user during his interaction with a media content delivery platform.

In such an environment, an intelligent personalisation system as envisioned within the LinkedTV project is called to face multiple benefits and challenges, including: understanding users' digital traces, aligning information from heterogeneous resources, exploiting freely available linked data, extracting user context, harvesting mass intelligence (i.e. learning and adapting group-specific knowledge) and bringing together the users' transactional and physical behaviour. On top of these issues, LinkedTV aims to reduce knowledge and user model dimensionality (i.e. explore methods for lightweight, compact, meaningful representation of user-pertinent information) in order to produce efficient content and concept filtering with regard to effectively handling resource and ethical constraints (i.e. scalable and privacy-preserving user modeling).

7   Acknowledgements

8   References

1. A. B. B. Martinez, J. J. Pazos Arias, A. F. Vilas, J. G. Duque, M. L. Nores, 2009. What's on tv tonight? An efficient and effective personalised recommender system of TV programs. Digest of Technical Papers International Conference on Consumer Electronics, 2009
2. 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).
3. Kay, J., 1994. Lies, damned lies and stereotypes: pragmatic approximations of users. Proceedings of the 4th Int.Conf. on User Modelling (UM'1994), Hyannis, USA, pp.175–184.

4. X. Amatriain, J. M. Pujol, and N. Oliver. 2009. I Like It... I Like It Not: Evaluating User Ratings Noise in Recommender Systems. In Proceedings of UMAP '09, Springer-Verlag, Berlin, Heidelberg, 247-258.

5. S. Dumais, E. Cutrell, JJ Cadiz, G. Jancke, R. Sarin, and D. C. Robbins. 2003. Stuff I've seen: a system for personal information retrieval and re-use. In Proceedings of SIGIR '03, ACM, New York, NY, USA, 72-79.

6. H. Ma, H. Yang, M. R. Lyu, and I. King. 2008. SoRec: social recommendation using probabilistic matrix factorization. In Proceedings of CIKM '08. ACM, New York, NY, USA, 931-940.

7. S. Yang, B. Long, A. Smola, N. Sadagopan, Z. Zheng, and H. Zha. 2011. Like like alike: joint friendship and interest propagation in social networks. In Proceedings of WWW '11. ACM, New York, NY, USA, 537-546.

8. 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.

9. 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).

10. 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

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

12. G. Adomavicius and A. Tuzhilin. 2008. Context-aware recommender systems. In Proceedings of RecSys '08. ACM, New York, NY, USA, 335-336.

13. C. Palmisano, A. Tuzhilin, M. Gorgoglione, 2008. Using Context to Improve Predictive Modeling of Customers in Personalisation Applications. IEEE Trans. Knowl. Data Eng. 20(11): 1535-1549.

14. T. Tsunoda, M. Hoshino 2008. Automatic metadata expansion and indirect collaborative filtering for TV program recommendation system. In: Multimedia Tools and applications, Vol. 36, (January 2008) , p. 37-54.

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

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

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

18. Vicon motion capture systems: http://www.vicon.com/products/viconmx.html

19. SoftKinetic company: http://www.softkinetic.com/

20. 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/

21. Hall, E.T. The Hidden Dimension. Doubleday, N.Y, 1966.

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

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

24. S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, and Z. Ives. 2007. DBPedia: a nucleus for a web of open data. In Proceedings of ISWC'07/ASWC'07. Springer-Verlag, Berlin, Heidelberg, 722-735.

25. 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.