NoTube – Making TV a Medium for Personalized Interaction

Lora Aroyo
Vrije Universiteit Amsterdam

Alex Conconi
TXT Polymedia

Stefan Dietze
Open University

Annelies Kaptein
Stoneroos Interactive Television

Lyndon Nixon
STI International

Christoph Nufer
Institut für Rundfunktechnik

Davide Palmisano
Asemantics

Luca Vignaroli
RAI Radiotelevisione Italiana

Milena Yankova
Ontotext Lab, Sirma Group Corp.

contact: l.m.aroyo@cs.vu.nl

ABSTRACT

In this paper, we introduce NoTube's vision on deploying semantics in interactive TV context in order to contextualize distributed applications and lift them to a new level of service that provides context-dependent and personalized selection of TV content. Additionally, lifting content consumption from a single-user activity to a community-based experience in a connected multi-device environment is central to the project. Main research questions relate to (1) *data integration and enrichment* - how to achieve unified and simple access to dynamic, growing and distributed multimedia content of diverse formats? (2) *user and context modeling* - what is an appropriate framework for context modeling, incorporating task-, domain and device-specific viewpoints? (3) *context-aware discovery of resources* - how could rather fuzzy matchmaking between potentially infinite contexts and available media resources be achieved? (4) *collaborative architecture for TV content personalization* - how can the combined information about data, context and user be put at disposal of both content providers and end-users in the view of creating extremely personalized services under controlled privacy and security policies? Thus, with the grand challenge in mind - to put the TV viewer back in the driver's seat – we focus on TV content as a medium for personalized interaction between people based on a *service architecture* that caters for a variety of content *metadata*, delivery *channels* and rendering *devices*.

Categories and Subject Descriptors

H.3.5 [**Information Storage and Retrieval**]: Online Information Services

General Terms

Design, Human Factors.

Keywords

Interactive Television, personalization, user modeling, context modeling, service oriented architecture, multimedia services.

1.INTRODUCTION

Back in 1960 – there was only one TV at home. It was bound to the living room where everybody would watch the same limited number of linear programs on a small set of channels. This classic TV notion is forever gone in 2010. The TV content has left the living room and has become dissociated from the TV tube. It has moved into the world of Internet and mobile technology. TV content is there in the train, at the back seat of the car, while sitting on a bench in the park. The available TV content, over the IP and broadcast channels, in terms of number and diversity has grown exponentially. The notion of channels has changed drastically: providers just offer a set of live and on-demand programs. 'Zapping until you drop' is replaced by services that enable you to generate your personalized schedule of TV programs. In terms of technology infrastructure there is no difference between TV and computer: the TV content and other streaming content live in a shared connected online world.

The advent of digital broadcasting technology for audio and video content has increased the amount of content from which a radio or TV user can choose dramatically. At the same time, people's preferences have become very much individualized, personal time constraints are complex, and the availability of digital storage technology at the recipient's side has opened a new degree of freedom. From this perspective context-awareness is an essential characteristic of an infrastructure that enables personalized TV service. Modeling and processing of contextual information about consumer preferences and behavior, about capabilities, performance and availability of devices and audiovisual streams is a core research field tackled within the NoTube.

Additionally the NoTube project will take into account community aspects (e.g., how one person can benefit from the media memories of others, or with whom they share common interests) in order to lift content consumption from a single-user activity to a community-based experience. These personalized and shared experiences typically take place in a connected multi-device environment. Also, the traditional of "device" in terms of media channels (e.g., telephone line, internet line, broadcast line) is rapidly disappearing, as the technological infrastructure is moving towards a single-cable concept. The role of these devices

then becomes what they actually should be, namely a vehicle for rendering in an optimal way content for a particular context.

In the landscape of TV content delivery we have to consider three types of parties:

- Broadcasters who provide live and/or on-demand programs.
- Telecommunications companies that bring TV content to consumers.
- TV viewers, who watch TV programs and who interact with the content and with other users of the same content.

The TV content data, the delivery channels and the consumer devices are of a varying nature and cannot be fixed in advance. For example, considerable variations exist in TV content metadata. Availability and popularity of delivery channels is very much dependent on the region. The consumer may use different rendering and interaction devices, such as a stationary TV with remote control, a handheld, or a laptop computer. Transparent development of those personalized services requires middleware that provides a layer of abstraction on top of these varying metadata formats, channels and devices. One of the key challenges of NoTube is therefore the development of a service architecture that can act as such a middleware. Figure 1 illustrates the NoTube vision of the complementary integration of Web, TV and mobile spaces by this middleware in order to enhance the TV experience.

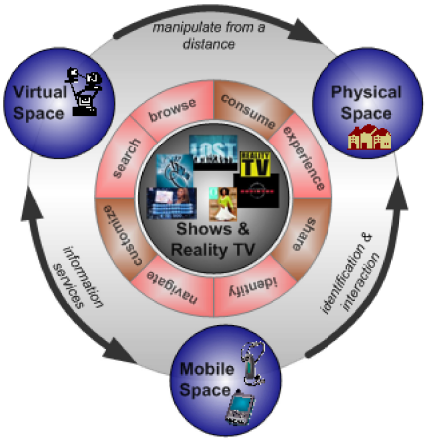

Figure 1: Complementary integration of Web, TV and Mobile spaces to enhance TV experience

In this paper, we introduce the core objectives of the project (Section 2), outline the envisioned service-oriented architecture (Section 3), describe the guiding use cases (Section 4), list the principle challenges (Section 5) and identify the target results (Section 6). In conclusion we discuss the contribution of NoTube to a new innovative system enhancing Interactive Television with advanced personalization technology.

2. CORE OBJECTIVES

We aim at a *user-centric approach* to investigate fundamental aspects of consumers' content-customisation needs, interaction requirements and entertainment wishes, which will shape the future of the "TV" in all its new forms. "New technology is transforming the TV industry", says Mark Thomson, BBC CEO for Observer. Watching TV more and more happens together with PC-related activities, e.g. chatting with friends, talking on the phone, searching on the Internet for related info about programs. Already in 2000, the trend was by "*couch-and-mouse toys* serving 27 million Americans who were able to watch TV and surf the Web at the same time – the "telewebbers". Nowadays, digital video recording software provides the facility to "time-shift" live television programming through a PC (e.g. via TiVo and ReplayTV). Companies are attempting to bundle "an electronic program guide into its software, along with personal viewing agents that can recommend broadcasted programs based on your viewing habits"[1]. For example, while watching "Dancing with the stars", you can vote for your favorite couple, read their biographies, find other programs this week where they appear, record your favorite dances and send them to friends; you can also join live chats during the program. However, "such interactive services stay TV-centric (in the device sense of the term TV) and use Internet and digital technology only to enhance the entertainment that you expect to see when you plop yourself down in front of the tube."[2]

Hence, our core objectives are developing semantics-based and context aware tools and services for (1) *personalized content selection*, (2) *packaging of content with additional information* (e.g. relevant Web info on program subjects), and (3) *social interaction in/consumer communities*. With semantic based we mean that we build on information integration and representation techniques developed within the Semantic Web area, enabling us to reason about related concepts in different information sources. Context-aware implies that the services are dependant on individual aspects of time, space, task and mood.

3. SERVICE-ORIENTED ARCHITECTURE

Service Oriented Architecture (SOA) is an emerging design model for distributed applications that clearly simplifies the development and re-arrangement of new and existing services. Applied technology for the realization of SOA are Web services, based on SOAP, UDDI and WSDL, that have the big advantage that they are platform independent and based on standard Internet protocols. To cope with the communication and coordination challenges implied by distributed systems of services a Semantic Web Services (SWS) framework is developed as a communication and coordination middleware that allows for decoupling applications in time, space, and reference. Moreover, due to implicit SWS technologies such as mediation, semi-structured data models and knowledge entailment, they also provide means to at least partly decouple applications in vocabulary, i.e. the terms and protocols applied in interactions. One of the primary aims of the NoTube project is to enhance the existing SWS technology to ensure a more adaptive, proactive and context-aware middleware built on top of established SWS frameworks such as WSMO.

To realize the vision of the core objectives we aim at a **service-oriented architecture** with four corner stones: (1) novel

[1] Control your TV with your PC, by Tom Springs, PC World, April 28, 2000

[2] Analysis: Interactive TV brings new tricks to an old box, by Tom Springs, PC World November 27, 2000

deployment of existing *methods for information integration* in the combined TV-Internet environment; (2) extensions of existing *user and context modeling techniques* to meet the demands of the distributed world; (3) development of *novel reasoning services* for personalized content recommendation; and (4) integration of community-oriented software and social interaction tools, such as chat channels.

Semantic Web Services enable the automatic discovery and selection of distributed resources - services and data exposed via Web services - for a particularly expressed user goal. Current results of SWS research are available in terms of reference ontologies, such as OWL-S and WSMO, as well as comprehensive frameworks (see DIP project results[3]). Particularly with respect to personalised delivery of media content, dynamic allocation of services, data and metadata is an important requirement to deliver the most appropriate media services to the end user. Context-adaptive delivery of broadcasting media has to consider two distinct context abstractions – i) *the provider context*, where broadcasts have to be directed to classified consumer categories, and ii) *the user and device context*, where media content is adapted to certain device and end-user criteria, such as supported resolution, formats or the preferred interactive feature add-ons. Hence, content adaptation tasks are usually carried out through a sequence of Web service invocations. In that, the *Semantic TV Resource Broker* will enable the dynamic discovery and orchestration of the most appropriate services for a given context.

4. DRIVING USE CASES

The work is steered by three **visionary use cases,** each addressing different dimensions of personalized TV-content interaction, including individual viewers and communities of viewers as well as multi-lingual and multimodal interaction:

- Personalized Semantic News, which focuses on the design and development of a system for the creation of a set of local personalized news services, able to acquire news items from generic broadcast streams, understand the meaning of video news items, understand the physical context in which news items are going to be shown, apply criteria for matching the user profile with the available news items.
- Personalized TV Guide with Adaptive Advertising, which focuses on context identification technologies (e.g. RFID/NFC tags, mobile phones, sensors), as well as different control interfaces (e.g. traditional TV remote control, Web-enabled remote control, mobile device actions) to offer a new Program Guide experience to the consumers; End-to-End Personalized Advertising creation, manipulation, delivery and consumption of personalized advertising messages, also allowing user-generated content and consumer participation to the content value chain.
- Internet TV in the Social Web, which focuses on next generation Web 2.0 technologies and social software to improve interactivity, participation and empowerment of TV consumers to become active protagonists of the TV value chain.

More generally, we will experiment with using media-based events in combination with semantics from the metadata in order to enable users to find content of a given type, with certain characteristics, which they or related users watched around the time of another media event. The reasoning infrastructure will help to overcome a number of limitations in current Electronic Program Guide (EPG) and move them to Personalized Program and Service Guides. Consider, for example, how to search for programs that are: (1) non-fiction, (2) produced between 1989 and 1995, (3) involve locations in Eastern Europe. Criterion 1 is resolvable using current EPGs. Criterion 2 is not generally supported by current EPGs, but is possible if the program information has been properly populated and indexed along a given timeline. If this is not the case, content analytics and data extraction techniques must be employed to generate the correct information. Criterion 3 again requires appropriate metadata, and also requires a sophisticated knowledge model that is able to represent and reason about part/whole relations between places, countries and regions (e.g. Sofia is in Bulgaria; Bulgaria is in Eastern Europe). The application of shared semantic spaces as medium for information integration of the multi-media system allows for combined data sources and reasoning within arbitrarily established communities.

We intend to improve the quality of News services available to end TV consumers not only for young generations, mobile and Internet surfers, but also for those categories of users who are fond of "traditional" sofa-TV (so-called "lazy" consumers) and not particularly ICT-literate (so-called "TV remote control" clickers). Instead of leaving to the final consumer all the decisions and all the personalization levers (e.g. accessing an EPG and programming by himself his/her own PVR), we propose the possibility to have a provider-consumer collaborative platform, so that part (or most) of the personalization activities are performed at providers' side on the basis on aggregated privacy-preserving semantic user-context-content profiles, allowing the most lazy users to placidly sit in the sofa and be able to access an highly personalized and interactive News service. Usually the TV consumer accesses to the entire broadcasted news program recorded, by using EPG information; this fact raises problems about news program EPG identification, about a huge memory space needed to record all the content and about the access time to a single news item of interest which may be very long.

5. RESEARCH CHALLENGES

The ICT landscape is developing into a highly-interactive distributed environment in which people interact with multiple devices (e.g. portable devices such as mobile phones and home equipment such as TV's) and multiple applications (e.g. computer programs such as Web browsers and dedicated Web services). The most recent proof for this is 'technology embraced by public' as Darren Waters calls it. The author shares the observations of Sean Wargo, director of industry analysis for the Consumer Electronics Association (CEA), at the Consumer Electronics Show 2007 in Las Vegas, one of the world's largest technology shows. People more than ever become early adopters of technology. Globally the industry is being driven by the shift away from old models - from physical to digital. New methods emerge for getting content such as TV programs via the Web. Almost half of the people want to watch TV content on their PC's. They want to make a bridge between a TV and a PC, perhaps even sitting in a home office. As we can see, technologies in these fields are rapidly progressing, but the user is lost. The

[3] DIP Project: http://dip.semanticweb.org

information overload is enormous and the content presented is hardly adapted to the prior knowledge, to the preferences and to the current situation of the user. Not only the users, but also the industry comes to the realization that content services and user experience are becoming crucially important.

Additionally, users are increasingly involved in multiple virtual environments (e.g. MySpace, Flickr, YouTube, Amazon, entertainment sites) in each of them with a different identity (e.g. login information, preferences). There is very limited integration of these user profiles, or if there exists integration it is not always under the control of the user and there is a lack of transparency in the use of personal data between different applications. As most of those services are relatively new and still aim at gaining critical mass of users, there is still not a methodological approach of how to assess the users experience and improve in an evolutionary way the provision of the services. Personalization in information retrieval and information presentation [3] has therefore become a key issue. Successful personalization experiments have been done, in e-commerce [1] and news websites [2], where the most common example is the Amazon.com recommendations. Personalization is seen as a key ingredient of the so-called "Web 2.0" applications [10,14]. However, such personalization is still local. In the example: the personalized information is only valid for Amazon; it cannot be used for other information services, nor can Amazon cater for different "modes" of a user, e.g. when looking for something personal, or when looking for a gift for his/ her kids or friends.

This new ICT landscape can be characterized with three key terms: Distributed Interactive, and Multi-device. In the project we use the acronym DIM to refer to this application setting. In DIM environments, individuals and groups make use of multiple devices in multiple distributed contexts (e.g. entertainment, personal information management). This results in complex interaction patterns requiring integrated views of distributed data collections, multiple modelling perspectives of content, user and environment data, as well as an increased need for personalized information presentation. To realize distributed personalization in DIM environments, we identify four main challenges: (1) *data integration and enrichment* - how to achieve unified and simple access to dynamic, growing and distributed multimedia content of diverse formats? (2) *user and context modeling* - what is an appropriate framework for context modeling, incorporating task-, domain and device-specific viewpoints? (3) *context-aware discovery of resources* - how could rather fuzzy matchmaking between potentially infinite contexts and available media resources be achieved? (4) *collaborative architecture for TV content personalization* - how can the information about data, context and user be put at disposal of both content providers and end-users in the view of creating extremely personalized services under controlled privacy and security policies?

To realize personalized TV in DIM environments we need to address the challenges outlined above. For the data integration problem there are already a number of general working solutions within the area of Semantic Web [11,12,13,6,7]. We plan to select appropriate solutions and deploy them in our project setting. Thus, data integration will not be the main focus of research in this project. User modelling is an area where multiple research examples, primarily in e-learning, have been presented [4,5]. However, the field still lacks concrete standardization of the user and context representation for both local and distributed applications [8]. Similar is the situation with personalized presentation of content, where the research is still in early stages [9]. The work here targets the use of open standards and reaching consensus for user and context modelling and personalized content presentation in DIM environments.

6. DIM ENVIRONMENTS

In the DIM environment we can perceive three "spaces", namely the physical, virtual and mobile spaces, see Figure 1. The user resides in the *physical space*, e.g., at home, on the street, in a museum. This space encompasses task-specific devices such as TV, VCR and information displays in museums. The user's prime goal in interacting with those devices is to consume and experience content. The rich-interaction environment of the Web plays the role of the *virtual space*, where users typically work with multiple domain-specific applications and perform information intensive tasks, such as searching, browsing and navigating in various content collections. The information prepared on the Web could be further used in the physical space: preferences for TV programs could be uploaded to a TV set-top box. The *mobile space* encompasses all mobile devices (iPod's, PDA's), each containing a fragmented portion of user preferences and data. In each of the spaces one can interact with content: one can physically look at a movie on the TV, view a trailer of it on a website or look at a picture sms'ed to your mobile phone. In earlier work we have identified six types of user interaction with content: search, browse, navigate, consume, experience, and share. We have also identified two additional user interactions, which operate on user data: namely identify and customize. All these interactions can be performed in each space. However, not all spaces provide the most efficient means for the user to realize the tasks easily. To enhance the user's experience and consumption of content in the physical world, we integrate the three spaces in such a way that the user can perform most efficiently the interactions in the dedicated spaces and use the results in the complementary ones. For example, identifying a user can more easily be done through a mobile phone. Browsing and searching is easier with a computer keyboard and mouse than with a TV remote control or a phone keypad.

7. CONCLUSION

In this paper we have presented the NoTube project and its vision for making TV a platform for personalised interaction. It will be possible to follow the activities and achievements of NoTube at the project website http://www.notube.tv

REFERENCES

[1] Ardissono, L., Console, L., and I. Torre. An adaptive system for the personalized access to news, 2001.

[2] Ardissono, L. and A. Goy. Tailoring the interaction with users in web stores. UMUAI, 10(4):251–303, 2000.

[3] Brafman, R.I., C. Domshlak, and S. E. Shimony. Qualitative decision making in adaptive presentation of structured information. ACM Trans. Inf. Syst., 22(4):503–539, 2004.

[4] Chin, D.N. Empirical evaluation of user models and user-adapted systems. UMUAI Journal, 11(1):181–194, 2003.

[5] Fink, J. and A. Kobsa. A review and analysis of commercial user modeling servers for personalization on the WWW, 2001.

[6] van Gendt, M., A. Isaac, L. van der Meij, and S. Schlobach. Semantic web techniques for multiple views on heterogeneous collections: a case study. In ECDL 2006, pp 426–437, 2006.

[7] Klein, M. Combining and relating ontologies: an analysis of problems and solutions. Workshop on Ontologies and Information Sharing, 2001.

[8] Kobsa, A. Generic user modeling systems. User Modeling and User-Adapted Interaction, 11(1-2):49–63, 2001.

[9] Light, M. and M. T. Maybury. Personalized multimedia information access. Commun. ACM, 45(5):54–59, 2002.

[10] Millard, D. and M. Ross. Web 2.0: hypertext by any other name? In HYPERTEXT '06, 27–30, 2006. ACM Press.

[11] Noy, N. and M. A. Musen. Prompt: Algorithm and tool for automated ontology merging and alignment. Workshop on Ontologies and Information Sharing, 63–70, 2001.

[12] Schreiber, G., et al, MultimediaN E-Culture Demonstrator. In The Semnantic Web – ISWC 2006, Athens, Georgia, volume 4273 of LNCS, pages 951–958. Springer Verlag, November 2006. Winner Semantic Web Challenge 2006.

[13] Shvaiko, P., and J. Euzenat. A survey of schema-based matching approaches. J. on Data Semantics, 146–171, 2005.

[14] Webster, D., et al. Context-orientated news filtering for web 2.0 and beyond. In WWW '06, 1001–1002, 2006. ACM Press.