Semantic integration of TV data and services: A survey on challenges, and approaches

Bassem Makni a, Stefan Dietze a and John Domingue a

a Knowledge Media Institute, The Open University Walton Hall, Milton Keynes, MK7 6AA, United Kingdom
E-mail: {b.makni,s.dietze,j.b.domingue}@open.ac.uk

Abstract. In this paper, we are surveying the impact of semantic Web and semantic Web services on enabling novel television features. These novel features include being Internet-based, mobile, interactive, personalized, social and semantic. Many research efforts have contributed to extending different aspects of television delivery and consumption, with respect to content production, metadata handling, semantic enrichment and recommendation. They adhere to the semantic Web vision for two goals: the seamless integration of their data and the automation of their Web services interoperation and composition. Mainly two semantic Web services approaches are used, namely a top-down and a bottom-up approach. We study the contribution of different Semantic Web and Semantic Web Services-based approaches to enable novel TV features.

Keywords: Next-generation TV, Semantic integration, TV data, TV services, WSMO, WSMO-lite, micro-WSMO, SAWSDL, semantic TV content annotation, social TV

1. Introduction

The television concept has evolved and ramified from its early form, of telegraph transmission of vision [38], to our contemporary perception of television. However, this contemporary perception is becoming increasingly vague, as today television content is scattered over broadcast streams, Web, and private Internet Protocol television (IPTV) networks, and is accessible, among classical sofa TV, via diverse enhanced devices such as smart phones, tablets, Apple TV1 and soon via Google TV2 devices. This content scattering perplexes the TV experience by increasing the time spent in searching for relevant media. Thus, novel TV features, including personalisation and social networks integration are required to enhance the TV experience.

Since many parties, from different backgrounds and with different concerns, are involved in materialising the novel TV features, for instance content producers, broadcasters, social scientists, interaction designers, etc. interoperability is a key issue. Chan and Zeng [22] impute the reason for the proliferation of metadata schemas to the requirements differences during the design phase with respect to intended users, subject domain, project needs, etc. These differences are radical among the TV parties, and are reflected at two levels:

- TV data integration
- TV Services interoperability

By TV data, we refer to the multimedia content and to the metadata describing this content. By TV services, we refer to the operations on the multimedia content and its metadata such as consumption, publishing and retrieval.

TV data integration challenge is to some extent similar to the integration of Web data, in terms of diversity and distribution. And thus, we conjecture that the lessons learnt from the Web of data movement are valuable for TV data integration, and especially defend the efficacy of semantic integration of TV data.
Similar challenges are posed with respect to the integration of services. The Semantic Web services (SWS) research efforts have been motivated with the need for automation of Web services related tasks such as discovery, orchestration, mediation, and composition. The first intention was the full automation of one or more of these tasks by providing conceptual models that comprehensively delineate the semantics of the services such as Ontology Web Language for Services (OWL-S) [71] and Web Services Modeling Ontology (WSMO) [94]. These efforts led to complex frameworks and tools for Semantic Web services annotation and brokering. Even with tools and communities built around these SWS approaches, their complexity hindered their uptake at a large scale for the automation of services tasks by the Service-Oriented Architecture (SOA) users. Moreover, the limited emergence of potential service automation scenarios, for instance, based on services which allow the creation of more complex orchestration or which provide analogous functionalities which could be exploited via SWS discovery, has put into question the actual potential for SWS automation.

Hence, efforts were made to redefine the SWS notion and the scope of its underlying technologies. More recently, a lightweight approach has been proposed with the aim of popularisation of SWS annotation use, namely with the standardisation of Semantic Annotations for WSDL and XML Schema (SAWSDL) [65] to ensure interoperability. While taking advantage from the lessons learnt from more heavyweight approaches, the lightweight approaches offer a less costly way of annotating services. However, they lack support for more complex reasoning, and therefore, provide only limited opportunities for the automation of services related tasks.

From our experience of using SWS to broker TV related services, within the European project NoTube that explores television’s future in the ubiquitous Web, two requirements for services management are prevalent: a) offering a lightweight means for service annotation and documentation, and b) enabling service brokerage through automating service discovery and orchestration [32].

The literature contains surveys on semantic integration [82,13,108,33], Semantic Web Services [93,35,86,73] and television evolution [2,12,63,41], but the impact of Semantic Web technologies on the television evolution has not yet been tackled as far as we know. Which is our main motivation for writing this survey as TV data is an interesting use case for Semantic Web technologies.

To clarify the specificity of TV related services, we start by introducing core features of next generation TV, implying different parties and services types (synchronous and asynchronous services, WS-* compliant and Web API etc.). Based on the requirements which arise from these aspects, we survey the main Semantic Web (SW) technologies usage within the TV domain and the SWS approaches that we classify as top-down or heavyweight and bottom-up or lightweight and their potential for supporting the above requirements. Finally, we foresee the upcoming research challenges in the SWS domain.

2. Next generation TV features

"New technology is transforming the TV industry", says Mark Thompson, BBC CEO for Observer [83]. We conjecture that the biggest catalyst of this transformation will be SW technologies.

In this section, we classify the new TV features and demonstrate in the following sections the impact of SW technologies to materialise these features.

2.1. The integration of the Internet and television

The integration of the Internet and television was being realised in both ways (a) by television services, like Video on Demand (VoD), becoming available over the Internet and (b) by television becoming connected to the Internet via connected TV and connected set-top boxes. Full integration with a built-in Internet browser in the TV set-top box will be popularised with projects such as Google TV and Apple TV. We focus on (a) way that we designate Internet based television.

Simpson and Greenfield [98] enumerate four ramifications; namely IPTV, Internet Protocol Video on Demand (IPVoD), Internet TV and Internet video; and define criteria that classify them. We note that these ramifications cross the television boundaries to video especially for Internet video and IPVoD. Which is consistent with the Telecommunication Standardization Sector of the International Telecommunication Union.

---

3http://www.notube.tv/
4http://www.google.com/tv/
5http://www.apple.com/appletv/
(ITU-T)\(^6\) definition of IPTV as “multimedia services such as television/video/audio/text/graphics/data delivered over IP based networks managed to provide the required level of QoS/QoE, security, interactivity, and reliability”.

Thus, our survey will cover Television and Digital video, as they share many research challenges for annotation and delivery.

2.1.1. Internet usefulness for TV experience

The two major advantages of Internet based television over broadcast television are (a) the built-in back channel and (b) the Internet Protocol (IP).

\(a\) The back channel or return path carries the user’s feedback to the broadcaster. This is compulsory for the new TV features such as interactivity and the personalisation.

\(b\) While digital video is a precisely timed and continuous stream and IP networks carry a loosely timed collection of data fragmented into discrete packets, both technologies are coupled for the following reasons:

- IP networking low cost owed to massive equipment production,
- IP standardisation
- IP independence from the physical communication layer, which could be either wired, wireless, 3G, or 4G based network.
- IP ubiquity, i.e. support by our quotidian devices, like mobiles, tablets, and console games, allows the TV mobility feature.

2.1.2. IP Video classification criteria

The main criteria used to classify IP video delivery systems are:

**Network type**  By network type we refer to network openness, while IPTV uses private networks to deliver content to subscribed users, IPVoD, Internet TV and Internet video are delivered via public networks typically the Internet.

**Quality of Service**  The Quality of Service (QoS) can be used to assign high priority to video packets so they are privileged by routers. However, this is meaningless in public networks such as the Internet where each application can mark its packets as high priority, so managed video delivery QoS is used only in IPTV [98].

Multicasting support  Multicast addressing\(^7\) is a network technology for the delivery of information to a group of destinations simultaneously using the most efficient strategy to deliver the messages over each link of the network only once, creating copies only when the links to the multiple destinations split. In the video delivery context, multicasting is crucial because of the data amount that could congest the network when redundant packets flood the routers [45]. IPVoD and Internet video support only unicasting to allow individual play functionalities such as pause and rewind. While Internet TV uses replicated unicasting, in which messages are sent one by one to each client [45], and IPTV supports multicasting.

**Delivery methods**  Two methods could be used for video delivery, either a HTTP based method called the progressive download or a streaming method via dedicated streaming protocols such as Real-time Transport Protocol (RTP) over User Datagram Protocol (UDP). The UDP stateless nature is useful for real time video streaming because dropping packets is preferable to waiting for delayed packets.

**Digital rights management**  Broadly refers to a set of policies, techniques, and tools that guide the proper use of digital content [102]. Video is one of the main applications of Digital rights management (DRM) especially in IPTV and IPVoD.

**Discussion**  We excerpt the criteria chosen by Simpson and Greenfield [98] to classify IP video in Table 1.

The novel technologies are continuously changing the way we consume and produce television content. Thus, internet video has been classified [98] and re-classified [99] during the last decade by refining the classification criteria to consider new technological advances. We foresee that the discussed classification will soon be made obsolete by initiatives like Hybrid Broadcast Broadband TV (HbbTV)\(^8\) and Google TV\(^9\).

The emergence of hybrid architectures, such as HbbTV, will solve the insufficiency of the existing networks infrastructure to deliver large video content, like High-definition television (HDTV) and 3D television (3D-TV).

Furthermore, Google TV predict a full integration of the Internet and television by providing a built-in

\(^6\)http://www.itu.int/ITU-T/index.html

\(^7\)http://en.wikipedia.org/wiki/Multicast

\(^8\)www.hbbtv.org

\(^9\)http://www.google.com/tv/
Table 1

<table>
<thead>
<tr>
<th>Criterion</th>
<th>IPTV</th>
<th>IPoD</th>
<th>Internet TV</th>
<th>Internet video</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network type</td>
<td>Private</td>
<td>Public</td>
<td>Public</td>
<td>Public</td>
</tr>
<tr>
<td>Quality of Service</td>
<td>ManagedQoS</td>
<td>UnmanagedQoS</td>
<td>UnmanagedQoS</td>
<td>UnmanagedQoS</td>
</tr>
<tr>
<td>Multicast support</td>
<td>Multicasting</td>
<td>Unicasting</td>
<td>Replicatedunicasting</td>
<td>Unicasting</td>
</tr>
<tr>
<td>Delivery method</td>
<td>RTP over UDP</td>
<td>Progressive download</td>
<td>HTTP streaming</td>
<td>HTTP streaming</td>
</tr>
<tr>
<td>Rights management</td>
<td>Strong with DRM</td>
<td>Strong often with DRM</td>
<td>Fairly strong</td>
<td>Weak or nonexistent</td>
</tr>
</tbody>
</table>

Internet browser within the TV, and by allowing cross-searching over the Internet and television content.

2.2. Interactive TV

Interactive Television (iTV) is an active watching experience engaging viewers in choices and actions. This aim of making the television more dynamic and participatory is as old as the television itself. However the last century attempts to make the TV more interactive were not followed by the expected uptake [54]. This was mainly due to the high costs [92] and the intrusive interfaces [11] making the interactivity cumbersome. Hence, Jensen [54] described the iTV as a vaporware10, which is an advertised product, often computer software, whose launch has not happened yet and might or might never happen. Recently new pragmatic approaches to iTV, consisting of building on the existing and lowering the challenges, produced a new airiness for iTV.

2.2.1. Interactivity types

Gawlinski [43] considers the lack of an agreed framework for describing different types of interactivity, as one of the iTV difficulties. However, the following taxonomy of TV interactivity types from Curry [28] is the most agreed one:

Distribution interactivity refers to controlling the content delivery but not the content itself.

Information interactivity consists of choosing the delivered information such as weather or local news.

Participation interactivity involves the viewer in actions and choices that bring dynamic content. A typical action is voting, and a possible choice is the camera angle during a soccer game.

2.2.2. Re-enabling the social aspect of TV

Initially, watching the TV was a social activity where the whole family gathered around their sofa TV to watch the news or the night movie and talk about it with friends the following day. Later on, with the increasing use of Personal video recorder (PVR), watching TV became an individual activity. That reduced the viewers discussions about programs, which, according to the water-cooler effect [31], these discussions could be more interesting and more entertaining for the viewers than the program itself.

Social TV is defined as opportunity to interlink people and provide communication features to create connectedness via the TV [56,21,70,109]. We define it as an adaptation of the social principle “It’s not what you know, it’s who you know” [79] to the TV context: “It’s not what you watch, it’s who you know watch”. And when we consider the asynchronous communication, from the taxonomy of TV sociability [25], it becomes “It’s not what you watch, it’s who you know watch or have watched”.

The study of social interactive television is also not new [21], as Wellens [111] already stated in 1979 that “interactive television represents means of linking individuals together by providing each with an electronically mediated representation of the other’s voice and visual presence”. However the lack of TV specific guidelines for interaction, enforced the use of Human computer interaction (HCI) techniques [44], which resulted in a cumbersome social interaction that does not meet the expected seamless TV experience.

Recently, reenabling the social aspect to the TV has gained augmenting interest, so that MIT Technology Review11 listed it in the ten most important emerging technologies of 2010. This interest is mainly owed to the new opportunities offered by virtual social networks and by the commercial potential of the social TV.

10 http://en.wikipedia.org/wiki/Vaporware
11 http://www.technologyreview.com/tr10/
We distinguish two means for social interactivity in the TV context:

1. Using ancillary devices: This explores the media multitasking practice where the user simultaneously uses Internet and mobile phones while watching TV. According to Nielsen Three Screen Report [81] survey about Television, Internet and Mobile Usage in the US, simultaneous usage rose in the first quarter of 2010 by 35% to reach 60% of TV viewers. Comcast’s Tunerfish\(^\text{12}\) uses a web and mobile interaction to allow friends from Twitter\(^\text{13}\) and Facebook\(^\text{14}\) to share feedback about TV shows. The ancillary device could also be an enhanced control such as a sensor-enhanced pillow [6].

2. Directly on the TV screen: Where social interaction is displayed on top of the watched program in form of avatars [26,80] for example.

2.3. Personalised television

Similarly to the Web content expansion phenomena, the TV evolution to Internet TV has been coupled with the scattering of TV and multimedia content, where the user struggles to find a relevant content, which perplexes the leisure time. Hence, the importance of the personalised feature within the next-generation TV.

The personalisation of TV could benefit from research advances in recommendation systems but requires previous adaptation to TV content and TV interaction. Ardissono et al. [3] enumerate the following challenges to enable personalized television:

**Viewer Modelling** The acquisition, representation, and utilization of information about viewers, such as their characteristics (e.g., gender and age), preferences, interests, beliefs, and their viewing behaviour. This includes individual and group modelling.

**Viewer Identification** The recognition of the TV viewer(s) to provide personalized services.

**Program Processing** Implying programs segmentation, summarisation, and indexing.

**Program Representation and Reasoning** Modelling the programs’ characteristics to measure similarities or dissimilarities between the different programs. Reasoning about programs can include a range of techniques, such as recommendation techniques based on collaborative filtering, for example.

**Presentation Generation and Tailoring** The selection, organization, and customisation of television material based on viewer queries, processed programs, and viewer models.

**Interaction Management** Adapt the human computer interaction techniques to the TV context. The human-TV interaction should include mechanisms for attention and dialogue management.

**Evaluation** of the user’s satisfaction with respect to speed and accuracy. The speed in which the system is adapted to the user’s preferences and accuracy in terms of precision and recall of the recommended programs.

2.4. Next-generation TV challenges

Obviously, materialising the previously discussed TV challenges involves many parties, from different backgrounds and with different concerns.

1. Network experts to adapt the Internet infrastructure for multimedia and TV data delivery.

2. HCI specialists to define TV interactivity patterns and human-TV interaction.

3. Personalisation experts to propose TV specific recommendation systems.

4. Social scientists to build social networks around the TV preferences.

And thus interoperability is a key issue for TV data integration.

Guenther and Radebaugh [47] define interoperability as “the ability of multiple systems with different hardware and software platforms, data structures, and interfaces to exchange data with minimal loss of content and functionality”. Since any loss of TV data content or functionalities implies degradation of reasoning capacities such as personalisation, the seamless interoperability of TV data is a high level requirement.

Chan and Zeng [22] defines three levels of interoperability:

**Schema level** when different schemas are used.

**Record level** the same schema is used with different semantic interpretations of the elements.

**Repository level** when accessing the data is dependant to the used repository, which hinders cross-collection searching.
The three levels of interoperability issues are omnipresent within the different TV parties.

In the following sections, we try to answer the following questions:

- To which extent can semantic Web technologies enable the semantic integration of TV data?
- How can SWS enable TV services interoperability?
- From our experience in brokering TV services, what hinders the SWS uptake?

3. Semantic integration of TV data

Ziegler and Dittrich [114] define the **semantic integration** as “the task of grouping, combining or completing data from different sources by taking into account explicit and precise data semantics in order to avoid that semantically incompatible data is structurally merged.”

Since TV data is a multimodal data composed of

- **Multimedia content**
- **Structured Metadata descriptions** of the multimedia content
- **Semi-structured metadata** with free text descriptions of the programs embedded in **Electronic program guide (EPG)** for example, the semantic integration of TV data benefits from research advances in each modality:
  - **Semantic integration** and retrieval of multimedia documents,
  - **multimedia metadata interoperability**,
  - **Natural Language Processing (NLP)** and semantic enrichment research.

The main efforts of applying semantic Web technologies for TV data integration aim a reasoning based personalisation of TV and bringing the social aspect to TV. Thus the next-generation TV challenges are:

1. Adapting the advances of semantic multimedia retrieval to the TV content.
2. Semantic integration of the different TV related metadata.
3. Semantic integration of semi-structured and structured TV data.
4. Reasoning based personalisation of TV content.
5. Enabling the Social TV.

We organise the following sections according to each challenge and discuss the impact of SWS to solve them.

3.1. Bridging the semantic gap

Effective management of multimedia assets, including content-based indexing and retrieval, impose a deep understanding of the content at the semantic level [23]. That could be performed either manually or automatically. On the one hand, manual semantic annotation of multimedia content suffers from subjectivity of descriptions, which hinders interoperability [58], and is far from being a scalable solution. On the other hand, the automatically extracted multimedia features are low-level perceptual features, faraway from the high-level semantic descriptions that match human cognition [58]. In order to improve **Content-Based Multimedia Indexing and Retrieval (CBMIR)** accuracy, the research efforts have shifted from designing sophisticated low-level features extraction algorithms to bridging this so-called **semantic gap** [67].

Kompatsiaris et al. [58] classify these efforts in the following categories: Relevance feedback [96], knowledge based [112] and multimedia ontologies. Liu et al. [67] distinguish a specific category for Web image retrieval, which is HTML text fusion with visual content from Web images, that we include into the multimodal fusion [9]. We focus on multimedia ontologies category as it is the most relevant for the Semantic Web domain.

Supported by the proved effectiveness of systems with limited context of application [112], the knowledge based approaches model the domain of application either explicitly or implicitly.

**Explicit** Model based approaches uses a priori domain-specific knowledge [1,103] for guiding low-level feature extraction, high-level descriptor derivation and symbolic inference [58]. Chang et al. [24] introduced the idea of semantic visual templates to link visual features to semantics, where each template represents a personalized view of concepts. Prior knowledge inspired from cine-
matic principles is relevant also in video classification [17]. Lighting level differentiate low light horror movies from well-lit comedies[17]. And motion speed separates fast action movies and sports from slow drama. Audio effects are also pertinent to automatically detect horror movies [76].

**Implicit** Uses machine learning techniques for discovering complex relationships and interdependencies between numerical image data and the perceptually higher level concepts [58].

Multimedia ontologies play a key role in modelling this knowledge in a shared formalisation that allows automatic bonding of high-level concepts from the model to the extracted low-level features.

### 3.1.1. Multimedia ontologies

Kompatsiaris et al. [58] motivates the usage of multimedia ontologies in formalizing the multimedia semantics as they fulfil the following requirements:

1. **Persistence**: Modelling the multimedia semantics evolution, such as the evolution of typical desk components, to allow usage in future applications.
2. **Consistency**: Precise and non ambiguous semantic annotations are crucial for efficient reasoning about the multimedia content.
3. **Context enabled**: As multimedia objects exist in context, modelling this context information is beneficial for multimedia retrieval [39].

We add that

4. syntactic annotations are liable to ambiguities and thus not interoperable nor interpretable by machines.
5. Semantic annotations refer to a knowledge formalised by an external ontology to help solving ambiguities via persistent and implicit annotations.
6. They are also operational annotations as they are intended to be consumed and generated by software agents.

Naphade et al. [78] have modelled large-scale concept ontology for multimedia (LSCOM) to enable automatic extraction of broadcast news video. Similar approaches to link low-level *Moving Picture Experts Group*-7 (MPEG-7) features to higher level concepts include [8,40,91,10]. Hauptmann et al. [50] have compared concept-based using LSCOM ontology and text-based retrieval accuracy over a collection of news videos and concluded the efficacy of concept-based approaches.

### 3.1.2. Linked multimedia

The success of the semantic Web vision has been limited to a small scope in enterprises and in virtual communities. This is essentially due to the scope of their domain knowledge that eases its modelling, added to the non maturity of ontologies merging algorithms. That led to isolated islets of semantic Web data. This ascertainment motivated the *Linked Open Data* (LOD) proposition [14]. This new vision defines the Semantic Web as “a technology for sharing data, just as the hypertext Web is for sharing documents” [15].

The linked data movement impulsed the adoption of the Semantic Web vision at a large scale by linking 25 billion *Resource Description Framework* (RDF) triples from 203 datasets (as of September 2010).

The need for this impulsion was also present within the semantic multimedia community, and hence the idea of adapting the linked data principles to the multimedia context. Burger and Hausenblas [18] enumerate the following principles to interlink multimedia data:

1. Follow the LOD principles
   - Use URIs as names for things
   - Use HTTP URIs so that people can look up those names.
   - When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL)
   - Include links to other URIs, so that they can discover more things.
2. Consider the contextual aspect to represent the semantics of multimedia content.
3. Deploy legacy multimedia metadata formats.
4. As the need to refer fragments of multimedia based on space and temporal parameters is fundamental [105], a mechanism to specify URIs for these fragments is mandatory.
5. Interlinking methods are essential in order to manually or (semi-) automatically interlink multimedia resources.

**Discussion**

The majority of the domains modelled by multimedia ontologies are relevant for TV content such as:

- News [78,51,100,55]
- Sports [113,30]
- Movies [29,90]
However the existing approaches for semantic annotation of multimedia are offline [7] due to the time consuming phases of features extraction and classification. That hinders the adoption for live programs broadcasting and reveals the need for real-time semantic annotation of multimedia content.

3.2. TV-related Metadata standards

We mentioned that digital video and TV share many concerns such as content annotation, and thus we will cover multimedia annotation standards namely MPEG-7 and Society of Motion Picture and Television Engineers (SMPTE) metadata dictionary and TV specific TV-Anytime and Public Broadcasting Metadata Dictionary (PBCore).

**MPEG-7** is a multimedia content description interface standardised by the International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC). The standard defines the MPEG-7 scope by addressing applications that can be stored (on-line or off-line) or streamed (e.g. broadcast, push models on the Internet), and can operate in both real-time and non real-time environments.

**SMPTE** The SMPTE Metadata Dictionary [95] is a large list of structured metadata elements grouped in the following classes: Identification, Administration, Interpretation, Parametric, Process, Relational, Spatio-temporal, Organisationally Registered Metadata, and Experimental Metadata. Although it was originally designed to be encoded in the Key-Length-Value (KLV) data encoding standard, an Extensible Markup Language (XML) serialisation is available.

**Standard Media Exchange Framework (SMEF)** The BBC has defined SMEF [16] to support and enable media asset management as an end-to-end process from commissioning to delivery to the home. The SMEF Data Model (SMEF-DM) provides a set of definitions for the information required in production, distribution, and management of media assets, expressed as a data dictionary and a set of Entity Relationship Diagrams.

**PBCore** The PBCore was created by the Corporation for Public Broadcasting (CPB) in the United States to provide a simple structure that its member stations and related communities can share. PBCore extends Dublin Core by adding a number of elements specific to audiovisual assets that falls into three groups:

1. Content: provides descriptive metadata elements.
2. Intellectual property: provides Rights management metadata.
3. Instantiation: contains all technical metadata about the physical or digital representation of the asset such as format, media type, duration etc.

**TV-Anytime** The TV-Anytime forum [17] is a worldwide project involving vendors, broadcasters, telecommunications companies, and the consumer electronics industry, which has defined an extensive bundle of specifications for the use of local storage at home in a specialized “set-top box” or in the TV set [27].

**Discussion** The coexistence of many metadata standards for TV is practically equivalent to the lack of standards as users will again fall to using non-interoperable metadata schemas. Which is inevitable within heterogeneous communities [22]. Thus the need for core multimedia ontologies to enable multimedia metadata schemas interoperability [53].

Moreover, the large number of elements of the discussed metadata schemas reveals their complexity, and the semantics of these elements remain implicit. For example, very different syntactic variations may be used in multimedia descriptions with the same intended semantics, while remaining valid MPEG-7 descriptions, which causes serious interoperability issues for multimedia processing and exchange [104]. Hence the need for multimedia ontologies unfolding metadata semantics and amending their interoperability at the records level.

**Multimedia ontologies for records interoperability**

Mai Chan and Lei Zeng [68] define the records level interoperability as the “efforts intended to integrate the metadata records through the mapping of the elements according to the semantic meanings of these elements”.

The multimedia ontologies used for records level interoperability include:

---


16 [www.bbc.co.uk/guidelines/smef](http://www.bbc.co.uk/guidelines/smef)

Core Ontology for Multimedia (COMM) [4] was designed by re-engineering the MPEG-7 standard in order to discover multimedia patterns. Patterns recognition was based on two of the main patterns of Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) which are Descriptions & Situations (D&S) and Ontology of Information Objects (OIO). The typical scenario “the decomposition of a media asset and the (semantic) annotation of its parts” reveals the two main functionalities of MPEG-7: decomposition and annotation. The decomposition consists of segmenting the multimedia content based on temporal, spatial or spatio-temporal descriptors. Then these segments are annotated with the MPEG-7 features descriptors. Following the D&S pattern, decomposition is a Situation (SegmentDecomposition) that satisfies a Description (SegmentationAlgorithm).

MPEG-7 Ontology Hunter [52] reverse-engineered a core subset of MPEG-7 specification to generate an RDF Schema (RDFS) ontology describing MPEG-7 elements semantics before generating a Web Ontology Language (OWL) version18.

Multimedia ontologies for schemas interoperability

The Multimedia Metadata Ontology (M3O) is a follow-up initiative to COMM based also on DOLCE but not restricted to MPEG-7 and thus capable of expressing all structural information of many multimedia metadata formats while preserving the abilities of the COMM.

3.3. Reasoning about TV data

The ultimate goal behind lifting the TV data to the semantic level is allowing reasoning about TV data in order to facilitate personalisation, recommendation and social TV features. Besides the semantic annotation of TV data, the viewers and their context should be represented at the semantic level to allow automatic matching.

3.3.1. Viewer profiling

Since user profiling and personalisation is a well established solution to the information overload problem [42], it stimulates viewer modelling activity to solve the similar problem of TV content overload. Gauch et al. [42] define user profiling as gathering and exploiting some information about users in order to be more effective. In ontology-based user profiling [74], the user profile is represented in terms of interesting concepts [46].

3.3.2. Personalised TV

The most used TV personalisation techniques are

- Content based: uses a metric to quantify the similarity between viewers’ profiles and programs based on their content description. Similarity estimation is time consuming [37] due to the TV content amount.
- Collaborative filtering: recommends programs based on estimated similar profiles. Despite its effectiveness is many domains, collaborative filtering for TV programs suffers from many issues [85] such as a) first-rater problem as new programs are not rated enough to be recommended, b) cold-start problem where new users did not rated programs yet and no valid recommendation could be suggested, c) sparsity problem consists on lack of overlap between two random viewers if they did not rated the same programs. O’Sullivan et al. [85] focus on c) as they consider it the most stringent problem of collaborative filtering for TV program recommendation.
- Social filtering: similarity between profiles is based on their friendship in social networks unlike estimation in collaborative filtering.

To overcome the shortcomings of each technique, TV recommender systems tend to use hybrid approaches [85,37,66].

Semantic personalisation of TV By semantic personalisation, we refer to both viewers’ profiles enrichment with semantics i.e. ontology-based viewer profiling and semantic representation of the TV content.

In SenSee framework [5], viewer profile and context are extended via ontologies describing time, geographical location and TV domain knowledge. The authors, Aroyo et al. [5], prove the advantage of ontology-based TV recommendation by drawing a quantitative comparison with free text approach.

Avatar [37] recommendation is based on hybrid approach using collaborative filtering and semantic similarity between the content and the user profile. The se-
mantic similarity is calculated according to a dedicated TV domain ontology\textsuperscript{19}.

The NoTube BeanCounter \textsuperscript{[106]} aggregates user’s activity from different social networks and uses the activity stream in TV programs recommendation. The aggregation from different Web sources illustrates one of the most important advantages of semantic technologies, for instance the NoTube BeanCounter aligns the data gathered from Last.fm with the BBC programmes ontology in SKOS (Simple Knowledge Organization System) \textsuperscript{[75]}.

3.4. Discussion

From the user’s perspective, the semantic integration of TV data brings the freedom of choice allowing the cross-collection searching over many Internet video and TV repositories while preserving a centralised profile adapted to each collection. That is materialised through the semantic interoperability between the different used multimedia metadata. However the content description is still generated manually or via crowdsourcing due to the remaining semantic gap between the computed multimedia features and the content concepts. Thus the semantic integration of TV data is dependant to the accuracy of the content description.

4. Semantic integration of TV services

Towards collaboration, the different next-generation parties adhere to the SOA to expose their services and consume others\textsuperscript{‘}. Specifically, the NoTube SOA is based on Web services to take advantage of the well-established Web architecture as a medium for services communication.

4.1. The Web of services

Mainly two approaches are used to adapt the Web to a Web of services: the first approach proposes a stack of specifications to support the services requirements such as communication, selection, security etc. giving this approach the Big Web Services name \textsuperscript{[88]}, the second approach, RESTful Web Services, claims that the success of Web is due to the maturity and ease of use of its Representational State Transfer (REST) architecture and that the same architecture could lead to the programmable Web. From the RESTful prospect, the WS-* specifications do not complement each other but usually overlap and compete, which is confusing Web Services designers. Moreover building Remote Procedure Call (RPC) upon Web is counter-intuitive and does not take advantages of the Web’s REST architecture. More objectively, Pautasso et al. \textsuperscript{[88]} made a qualitative technical comparison based on architectural principles and decisions and concludes that REST is well suited for basic, ad-hoc integration scenarios and that WS-* is more flexible and addresses advanced quality of service requirements.

4.2. TV services

It became common practice throughout the last decade; to expose all sorts of multimedia content and metadata stored in one particular repository through a set of Web APIs. The motivation for this practice is to allow the aggregation of the multimedia content and its tailoring to users’ requirements. Which delegates the adaptation of multimedia consumption to third parties developers and exempt the multimedia providers from maintaining different platforms end-points such as Android\textsuperscript{21} and iOS\textsuperscript{22} platforms.

To name but a few, major TV broadcasters such as BBC expose their data via Web endpoints\textsuperscript{23} and provide APIs to process these data, and YouTube\textsuperscript{24} data API\textsuperscript{25} allows a program to search for videos, retrieve standard feeds, and see related content. However, these APIs are not standardised in terms of inputs, outputs, nor invocation methods, and specific clients should be designed to each API. That raises the repository level interoperability issue. In the following section, we present the major Semantic Web services approaches that tackle this issue generally and in TV context specifically.

4.3. Semantic Web Services

Whatever the used approach RESTful or WS-* Web Services, performing complex operations such as services selection, composition, or mediation requires hu-

\textsuperscript{19}http://avatar.det.uvigo.es/index-i.html
\textsuperscript{20}http://purl.org/ontology/po/
\textsuperscript{21}http://www.android.com/
\textsuperscript{22}http://www.apple.com/iphone/ios4/
\textsuperscript{23}http://backstage.bbc.co.uk/data/Data
\textsuperscript{24}YouTube.com
\textsuperscript{25}http://code.google.com/apis/youtube/getting_started.html\#data_api
man intervention. However, these operations should be automated as the users required functionalities are seldom achieved via one Web Service; hence the need for automated services composition and mediation in large scale context. Such as the Web 1.0, the software agents could not reason about Web Services as they see them as inputs and outputs without any conscience if, for example, the received message contains a ranked list of programs from a recommendation service or a program description from an EPG service. And to automate services reasoning, the software agents should process the services messages and operations at the semantic level via shared ontologies. Which is the research field of SWS [72] to solve at least one of the following challenges.

4.3.1. Discovery
Semantic Web Services discovery consists on retrieving the relevant services that achieve the requested functionalities. Mainly discovery is based on advertising service capabilities in a centralised or distributed registry, and matchmaking these capabilities with the requests.

4.3.2. Composition
We noted that complex functionalities are rarely achieved via one Web Service, yet many functionalities from different Web Services should be combined to achieve complex requests. Most of the proposed approaches for automatic Web Services composition are inspired by the researches in cross-enterprise workflow and AI planning [93].

4.3.3. Mediation
Composing different Web Services from different providers and which are not initially designed to cooperate raises another challenge which is Web Services mediation. Mediation aims to adapt services outputs to be consumed by the following service(s) in the composition chain.

4.3.4. Choreography
The web Services choreography handles the interaction between the services invoked in the composition, from a global point of view.

4.4. Preliminaries
In this section we define required notions to compare the different SWS approaches.

4.4.1. Capabilities representation
There are two approaches to represent the capabilities of a Web Service, the first one is based on an extensive ontology of functions where capabilities advertisement is done by binding to a class of homogeneous functions within a taxonomy of services such as flight booking, or transportation service. The second approach models the flow of state transformations, in occurrence the flight booking service requests a departure and arrival cities, a departure and arrival dates and a credit card, and changes the state by decreasing the number of available seats and withdrawing the due amount of money.

4.4.2. Artificial intelligence planning
Stuart et al. [101] define artificial intelligence planning (AI planning) as “a kind of problem solving, where an agent uses its beliefs about available actions and their consequences, in order to identify a solution over an abstract set of possible plans.” This agent, the planner, accepts three inputs

1. a formalised description of the initial state of the world.
2. a formalised description of the agent’s goal (the desired behaviour).
3. a formalised description of the possible actions that can be performed (the domain theory)

and outputs a sequence of actions that, when executed in any world satisfying the initial state description, will achieve the goal. [110]

4.4.3. Cross-organisational workflow
A workflow is an abstraction of a business process. It comprises a number of logic steps (known as tasks or activities), dependencies among tasks, routing rules, and participants [20]. When many organisations intervene in the business process, its abstraction is a cross-organisational workflow that could be assimilated to Web Services composition.

4.4.4. Abstract State Machines
The Church-Turing thesis states that any real-world computation can be translated into an equivalent computation involving a Turing machine. However, steps number required by the machine is not bounded and even simple operations could be simulated by a large number of steps. So the importance of the Abstract State Machine (ASM), introduced by Gurevich [48] as every sequential [49] or parallel algorithm [16] is behaviorally equivalent to a correspondent ASM. The behaviour emulation is a set of transitions between abstract states where
4.5. Semantic Web Services approaches

We distinguish two approaches that aim to solve one or more of the semantic Web Services challenges, which are the top down and bottom up approaches.

4.5.1. Top down approaches

Top down approaches provide conceptual frameworks and languages to describe the semantics of Web Services before grounding these descriptions to the services.

The OWL-S approach The OWL-S aims to provide building blocks for encoding rich semantic service descriptions that builds naturally upon OWL [71]. The OWL-S approach consists on an Upper Ontology for Services with three interrelated sub-ontologies: Fig. 1

Profile ontology for describing the service functionalities in order to advertise the service and match it with the requests.

Process model ontology for behavioural description with the intention of service invocation, enactment, composition, monitoring and recovery.

Grounding ontology bonds the process model with detailed specifications of the service from Web Services Description Language (WSDL).

Due to this enrichment of expressiveness, OWL-S extends the WSDL operations to a more abstract construct “atomic process” that extends inputs and outputs and introduces preconditions and effects (IOPE). The inputs and outputs are typed following the OWL typing system which allows binding to a class hierarch. Grounding these types to WSDL and vice-versa could be defined via Extensible Stylesheet Language Transformations (XSLT). Preconditions and effects are defined via rule languages such as Rule Markup Language (RuleML) or OWL Rules Language (ORL) to define the state respectively before and after the invocation of the service. We summarise how the OWL-S approach tries to solve the semantic Web Services challenges:

**Services discovery with OWL-S** For capabilities description, OWL-S supports both services taxonomy and state transformation approaches. Services taxonomy is formalised by sub-classing the Services Profiles which is also used to define the state transformations. Then to match the capabilities with the request, either by searching a subsumption relation between the requested service class and a class from the services taxonomy, or by matching both the inputs and outputs of the request and the advertised services also via subsumption relations. Fig. 2 illustrates the OWL-S discovery approach.

**Services composition with OWL-S** OWL-S based composition naturally falls into the AI planning section, as OWL-S provides

1. the formalised description of the initial state of the world, in terms of preconditions.
2. the formalised description of the desired goal in terms of users required outputs and effects.
3. the domain theory by modelling the OWL-S atomic process as an action that transforms the inputs into outputs.

The next step is the choice of a suitable AI planning algorithm for Web Services composition. Oh et al. [84] draw a decision tree in Fig 3 that pilots this choice according to the scale of the available services and the complexity of composition. A simple composition involves a sequential AND operator while a complex composition is expressed by AND, OR, XOR, NOT operators and by constraints.

**The WSMO approach** The WSMO [94] approach reuses the main concepts identified in the Web Service Modeling Framework (WSMF) [36] to define Semantic Web Services:

**Ontologies** provide the terminology used by other WSMO elements to describe the relevant aspects of the domains of discourse.

**Web services** represent computational entities able to provide access to services that provide some value in a domain. The terminology defined by
the ontologies is used to describe the Web services capabilities, interfaces, and internal working.

**Goals** model the user view in the Web service usage process in terms of requested functionalities.

**Mediators** handle interoperability problems between different WSMO elements, at the data level to resolve mismatches between different used terminologies, at the protocol level to ensure communication between Web services and on process level when combining Web Services.
Moreover all these concepts could have Non-Functional properties.

Web Services mediation with WSMO WSMO supports SWS mediation naturally via the Mediators concept. Indeed, WSMO defines four types of mediators of two categories Fig. 4:

**Refiners** express the refinement relation between elements

- **OO-Mediators** import ontologies and resolve possible representation mismatches between them.
- **GG-Mediators** express goals refinement and equivalence.

**Bridges** enable heterogeneous elements interoperation.

- **WG-Mediators** express total or partial fulfillment of desired goals by exposed Web Services.
- **WW-Mediators** deals with heterogeneity problems between Web Services that could appear during composition and orchestration tasks.

Web Services choreography with WSMO WSMO choreography inherits the core principles of ASM, namely state-based, represents a state by a signature, models state changes by transition rules[97]. The state signature is expressed in terms of instances of concepts or relations from a state ontology. The state change consists of new instance or new value for the relation attribute which leads to the notion of evolving ontologies by analogy to the evolving algebra, the first name of ASM.

4.5.2. Bottom-up approaches

The bottom-up approach builds incrementally upon existing Web services standards [65].

**The SAWSDL approach** SAWSDL recommendation [65] forms the first brick upon WSDL that gears up semantic annotations of services by providing Model Reference and Schemas Mapping extensions [60]. Where

**Model Reference** is an extension attribute, sawsdl:modelReference, applicable to any WSDL or XML Schema element to point to one or more semantic concepts, in order to describe the meaning of data or to specify the function of a Web service operation.

**Schemas Mapping** consists of transforming data from XML message format to a semantic model and vice-versa. The former transformation is called lifting and expressed by the sawsdl:liftingSchemaMapping attribute and the former is called lowering expressed by the sawsdl:loweringSchemaMapping attribute.

**The WSMO-Lite approach** As SAWSDL in itself does not define the semantics of Web services, but offers means to link the WSDL descriptions to ontologies, the Web Services Modeling Ontology Lite (WSMO-Lite) approach [107] defines a service ontology that bonds semantics to the service description via SAWSDL attributes. These semantics are expressed by the following WSMO-Lite ontology concepts:

- **Ontology** a subclass of owl:Ontology that defines a container for a collection of assertions about the information model of a service.
- **ClassificationRoot** defines a root class for a taxonomy of services functionalities.
- **NonFunctionalParameter** allows the description of domain specific nonfunctional properties.
- **Axiom** sub-classed in Condition and Effect to form a service capability.

4.5.3. Semantic annotation of RESTful services

Since there is no agreed machine-processable description language for RESTful Web services [69], the MicroWSMO [59] approach for semantic RESTful services builds on top of hREST [62], a microformat [57] that enables the creation of machine-processable descriptions on top of existing HTML descriptions. MicroWSMO tends to provide a SAWSDL-like annotation of Restful services that could be bonded to WSMO-Lite ontology in order to provide RESTful and WSDL based services interoperability. The Figure 5 illustrates this relative positioning.
4.6. Semantic TV services

Our experience of brokering TV-related services within the NoTube project offered us a fertile ground to apply, compare, and adapt different semantic Web services approaches [32]. We introduce the specific NoTube challenges, the different semantic TV services approaches, and their shortcomings.

4.6.1. NoTube use-case

In order to illustrate the challenges with respect to service-related tasks, we describe one of the main use cases driven by the TV broadcast industry partners within the NoTube project - namely, the requirement to provide personalized content and metadata delivery to users. Here, the basic feature is the matching of heterogeneous users’ profiles, e.g. including interests, preferences, and activity data, and user contexts (e.g. current location and viewing device) to filter and deliver TV content from a variety of sources. Addressing this particular use case in a service-oriented manner involves selecting, and orchestrating between numerous services that provide various functionality, for instance, to aggregate users’ topic interests based on their social networking activities, retrieve EPG data from various sources, and provide recommendations based on a dedicated algorithm. To support the highly service-oriented nature of the project, two major goals need to be supported: a) support of distributed developers with lightweight service annotations, and b) support of application automation with Semantic Web Service brokerage. In the early stage we focused on b) via a top down approach, namely, IRS-III [34]. However, the later need for the distribution of services annotation revealed the need for lightweight approach via iServe [89].

4.6.2. The IRS-III framework

IRS-III is a semantic execution environment that adopts the WSMO approach, videlicet that a service description is expressed in terms of Goals, Mediators, Web Services and Ontologies. These are described in a formal representation language, for instance, OCML [77]. IRS-III supports capability-based invocation: the request is a goal to be achieved via the following intermediated operations [19]

1. discover potentially relevant Web services;
2. select the set of Web services which best fit the incoming request;
3. mediate any mismatches at the conceptual level;
4. invoke the selected Web services whilst adhering to any data, control flow and Web service invocation constraints.

Given that IRS-III directly aims at automating service execution related aspects, the interface covers choreography and orchestration descriptions. Choreography addresses the communication between the IRS-III broker and a Web service, and is described as so-called grounding. The IRS-III grounding mechanism supports REST-based, SOAP-based, and XML-RPC based services [64]. Grounding involves two processes referred to as lifting and lowering. Lowering involves transforming input parameters at the semantic level to data input to the service at the syntactic level. Lifting involves the opposite transformation, i.e. transforming the data output from the service at the syntactic level into an ontological object at the semantic level.

At the semantic level the orchestration is represented by a workflow model expressed in OCML that describes the flow of control between Web services. The IRS-III orchestration model supports the main control-flow primitives of sequence, selection, and repetition.

4.6.3. The iServe Linked Services approach

iServe supports publishing service annotations as linked data - Linked Services - expressed in terms of a simple conceptual model that is suitable for both human and machine consumption and abstracts from existing heterogeneity around service kinds and annotation formalisms. Particularly iServe provides:

- Import of service annotations in a range of formalisms (e.g., SAWSDL, WSMO-Lite, MicroWSMO, OWL-S) covering both WSDL services and Web APIs;
- Means for publishing semantic annotations of services which are automatically assigned a resolvable HTTP URI;
– Support for content negotiation so that service annotations can be returned in plain HTML or in RDF for direct machine consumption;
– SPARQL endpoint allowing querying over the services annotations;
– REST API to allow remote applications to consume and provide annotations.
– Support for linking service annotations to existing vocabularies on the Web.

In order to cater for interoperability, iServe uses what can be considered the maximum common denominator between existing SWS formalisms, the Minimal Service Model (MSM). The MSM, first introduced together with WSMO-Lite and hRESTS [19], is thus, a simple RDFS ontology able to capture (part of) the semantics of both Web services and Web APIs in a common model. MSM is extensible to benefit from the added expressivity of other formalisms. The MSM, denoted with the 'msm' namespace in Fig. 6, defines Services as having a number of Operations each of which has an Input, Output MessageContent, and Faults. In turn, a MessageContent may be composed of MessageParts which may be mandatory or optional. iServe additionally uses the SAWSDL, WSMO-Lite and hRESTS vocabularies.

4.6.4. IRS-III and iServe comparison

We mentioned that our migration to iServe based annotation for TV services was motivated by the need for lightweight approach that allows distributed Web services developers to maintain their annotations. In table 3, we outline a comparison between the prerequisites for services annotators to master each framework.

Table 3

<table>
<thead>
<tr>
<th>IRS-III and iServe prerequisites comparison</th>
<th>IRS-III</th>
<th>iServe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Programming languages</td>
<td>LISP</td>
<td>OCML</td>
</tr>
<tr>
<td>Conceptual model</td>
<td>WSMO</td>
<td>MSM</td>
</tr>
</tbody>
</table>

In figure 6, we detail the conceptual comparison between IRS-III and iServe. That reveals the closeness between MSM concepts and Web services notions such as msm:Method and msm:MessageContent, unlike more abstract WSMO concepts such as wm:Goal and wm:Mediator. Which implies bigger effort from the annotators to master the WSMO conceptual model before being productive.

Listings 1 and 2 excerpt respectively annotations of Lupedia, a semantic enrichment service, within iServe and IRS-III. Ostensibly, the iServe annotation is more compact, however it cannot cover the lifting and lowering operations, as in IRS-III annotation. We recall that lowering and lifting operations are mandatory for execution support.

While allowing handy annotation of services, iServe lacks the support for complex Web service tasks such as choreography and mediation (Table 4). Dietze et al. [32] express the need for both facets: handy annotation and complex reasoning support for TV services. And propose a two-stage service annotation approach that links iServe and IRS-III annotation via cross-referencing.

Table 4

<table>
<thead>
<tr>
<th>IRS-III and iServe capabilities comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Execution support</td>
</tr>
<tr>
<td>Mediation</td>
</tr>
<tr>
<td>Choreography</td>
</tr>
<tr>
<td>Annotations publishable as linked data</td>
</tr>
</tbody>
</table>

5. Conclusions

The TV and multimedia are fertile and challenging domains for Semantic technologies application. The fertility arises from the heterogeneity of the different actors, and the challenges are centred around the TV and Web data differences. Certainly, the quantitative difference is considerable but more important, is data nature: the raw TV data are not human-friendly as opposite to Web data, mainly generated for Human consumption. For instance, we can imagine manual summarisation of a Web document by summarising the first paragraph of each section. However, the topical structure of a multimedia document is hidden. The impact of automatic multimedia management is hindered by the so-called semantic-gap.

Thus, the semantic integration of TV multimedia could be biased by the shortcomings of automatic multimedia annotation algorithms and at the same time cannot benefit from crowd-sourcing approaches to generate fined multimedia descriptions.

At the metadata level, the semantic interoperability via shared ontologies tends to be the remedy to
the plethora of multimedia metadata schemas. That allowed cross-collection searching and recommendation from different programs providers.

Besides TV data interesting application of semantic technologies, we considered the TV services inte-
Listing 2: WSMO/OCML-code of LUPedia service

```ocaml
(DEF-CLASS GET-LUPEDIA-ENTITIES-GOAL (GOAL)
  ((HAS-INPUT-ROLE :VALUE has-text)
   (HAS-OUTPUT-ROLE :VALUE has-lupedia-entities)
   (has-text :TYPE String)
   (has-lupedia-entities :TYPE List)))

(DEF-CLASS GET-LUPEDIA-ENTITIES-WS-PUBLISHER-INFORMATION (PUBLISHER-INFORMATION)
  ((HAS-WEB-SERVICE-HOST :VALUE "lupedia.ontotext.com")
   (HAS-WEB-SERVICE-LOCATION :VALUE "/lookup/text2xml")))

(DEF-RULE LOWER-FOR-GET-LUPEDIA-ENTITIES-GOAL
)
(DEF-RULE LIFT-FOR-GET-LUPEDIA-ENTITIES-GOAL
)
(extract-lupedia-entities-from-xml ?xml ?list-of-lupedia-entities
if (= ?list-of-lupedia-entities 
  (setofall ?lupedia-entity
    (and 
      (?xml:tag ?lookupsEl ?rootContents)
      (?xml:contents ?lookupsEl ?lookupsContents)
      (?xml:tag ?instanceURIEl ?lookupsContents)
      (?xml:tag ?classURIEl ?lookupsContents)
      (?xml:value ?lupedia-entity (#_LUPediaEntity
        ?instance-uri
        ?class-uri))))))
```

In light of these conclusions, we foresee the next research focus on building on top of SWS lightweight approaches to support execution, choreography, composition, and mediation. The real challenge consists of maintaining the lightweight aspect when incrementally augmenting these approaches, as it is the key for large scale uptake.

**References**


[8] B. Makni et al. / Semantic integration of TV data and services 19


[50] S. Kepplinger, F. Forster, J. Igelbock, and M. Obrist. Strengthen the TV Experience beyond the TV.


N.F. Noy. Semantic integration: a survey of ontology-based
B. Medjahed, A. Bouguettaya, and A.K. Elmagarmid. Com-
S. McIlraith, T.C. Son, and H. Zeng. Semantic web services.
M. Nathan, C. Harrison, S. Yarosh, L. Terveen, L. Stead,
B. Nardi, S. Whittaker, and H. Schwarz. It’s not what you
M. Naphade, JR Smith, J. Tesic, S.F. Chang, W. Hsu,
E. Motta. An overview of the OCML modelling language. In
D. Martin, M. Paolucci, S. McIlraith, M. Burstein, D. Mc-
Dermott, D. McGuinness, B. Parsia, T. Payne, M. Sabou,
M. Solanki, et al. Bringing semantics to web services: The
OWL-S approach. Lecture Notes in Computer Science, 3387:
26–42, 2005.
S. McIlraith, T.C. Son, and H. Zeng. Semantic web services.
B. Medjahed, A. Bouguettaya, and A.K. Elmagarmid. Com-
posing web services on the semantic web. The VLDB Journal-
The International Journal on Very Large Data Bases, 12(4):
351, 2003. ISSN 1066-8888.
S.E. Middleton, N.R. Shadbolt, and D.C. De Roure. Onto-
logical user profiling in recommender systems. ACM Trans-
ISSN 1046-8188.
A. Miles and S. Bechhofer. SKOS simple knowledge organi-
S. Moncrieff, S. Venkatesh, and C. Dorai. Horror film genre
typing and scene labeling via audio analysis. In Multimedia
E. Motta. An overview of the OCML modelling language.
In the 5th Workshop on Methods and Languages. Citeeseer, 1998.
M. Naphade, JR Smith, J. Tesic, S.F. Chang, W. Hsu,
L. Kennedy, A. Hauptmann, and J. Curtis. Large-scale con-
cept ontology for multimedia. Multimedia, IEEE, 13(3):86–
91, 2006. ISSN 1070-986X.
B. Nardi, S. Whittaker, and H. Schwarz. It’s not what you
M. Nathan, C. Harrison, S. Yarosh, L. Terveen, L. Stead,
and B. Amento. CollaboraTV: making television viewing so-
cial again. In Proceeding of the 1st international conference
on Designing interactive user experiences for TV and video,
watch: Technology enhances the video experience. Tech-
tical report, Nielsen, 1st Quarter 2010. URL http:
//en-us.nielsen.com/content/nielsen/en_us/insights/nielsen_a2m2_three.html.
N.F. Noy. Semantic integration: a survey of ontology-based
0163-5808.
The Observer. We must change again, says white city’s
radical, 2006. URL http://www.guardian.co.uk/
business/2006/jan/08/bbc.broadcasting.
Seog-Chan Oh, Dongwon Lee, and Soundar R. T. Kumara.
A comparative illustration of ai planning-based web services
//doi.acm.org/10.1145/1124566.1124568.
D. O’Sullivan, B. Smyth, D.C. Wilson, K. Mcdonald, and
A. Smeaton. Improving the quality of the personalized elec-
tronic program guide. User Modeling and User-Adapted In-