Linking Web Content Seamlessly with Broadcast Television: Issues and Lessons Learned

Jan Thomsen[1], Raphaël Troncy[2] and Lyndon Nixon[3]

[1] Condat AG, Berlin, Germany
<jan.thomsen@condat.de>
[2] EURECOM, Sophia Antipolis, France
<raphael.troncy@eurecom.fr>
[3] MODUL University, Vienna, Austria
<lyndon.nixon@modul.ac.at>

Abstract. The LinkedTV[4] platform integrates all necessary functionality for linking television and the Web into a managed workflow for media owners. This position paper reflects the issues we experience in implementing **Linked Television** using web specifications as well as our insight regarding applying the same ideas within the broadcast industry.

Keywords: LinkedTV, broadcast, media fragments, media Annotations, hypervideo, second screen

1 Linked Television: a Concept for Future Web-TV interlinking

As Television and the Web converge – television content is transmitted digitally, possibly over IP, to devices which are equally Web capable, while Web content includes audiovisual streams consumed over connected devices in the same manner as classical television – we observe that the actual meaning of "television" changes. On a connected device, (broadcast) television content, other audiovisual content and other Web content may be consumed in parallel or sequence, and the fact that something is "television" or "Web" blurs in the consumers' awareness. Is YouTube playing on a SmartTV now television? Is a catchup TV stream on the laptop still television? Linked Television is a vision of future television consumption, where TV and Web content is seamlessly interlinked, for example while watching the news the viewer is getting background information on the stories at their fingertips. Similarly, it is seeing a painting in a TV program and being able to identify the artist and the museum where it hangs. It is facilitating what the Second Screen generation is already doing, but now bringing this within the same digital ecosystem, so that content owners retain the connection with the viewer and control about the information being linked with their media.

[4] http://www.linkedtv.eu/

The concept of Linked Television is made up of an initial audiovisual stream, i.e. a TV program (termed the *seed content*), a set of concepts determined to be relevant to parts of that audiovisual stream (termed the *annotation*), and a set of digital contents which are associated to each concept as being informative or illustrative of that concept (termed the *enrichment*). Within a Linked Television environment, the viewer of supported seed content would be able to access a LinkedTV application on some connected device able to recognize which content was being consumed, retrieve the associated annotations and enrichments, and make them accessible to the viewer. This concept does not impose any expectations regarding which device runs the application nor on which screens the seed content, annotations or enrichments would be displayed, recognising that the future viewing environment may involve several devices capable to connect to the Internet and to share content between them.

This triple layered approach in Linked Television allows first that an abstract representation of the perceived topics and objects of relevance in the TV program is first captured, allowing for possible subsequent personalization to the current viewer and his/her particular interests or needs in the current viewing context. Then, enrichments are determined for that TV program based on the accepted topics and objects of relevance in the TV program, allowing that the experience of the enriched TV program could differ from viewer to viewer or from context to context, as opposed to current second screen companion applications whose associated content is hard linked to the running TV program in advance.

2 The Linked Television implementation

The Linked Television concept is currently being implemented on the Web, outside of the traditional TV broadcast networks and technologies. There are various reasons for this related to the greater availability of suitable specifications, and promise of interoperability across devices and applications. In this case, the seed content is an Internet audiovisual stream referenceable via a HTTP URL and the focus is on a recorded resource which should exhibit a persistent Web address. The ability to annotate parts of the seed content pre-supposes a model for decomposing the audiovisual stream into its parts, which could be temporal or spatial. While different annotation models provide descriptive elements to model such decompositions, e.g. MPEG-7, subsequent references back to specific parts become complex and difficult to support in different applications (e.g. the use of XML in MPEG-7 makes XPath necessary to refer to a `Decomposition`, which in turn requires applications to run XML parsers and support XPath syntax). In comparison, the W3C Media Fragment URI Specification defines a simple syntax to refer to a spatial and/or temporal fragment of some online media as a suffix to its URL (using the fragment part of a http URI).

Conceptual annotation relies also heavily in Linked Television on Web annotation models. The Semantic Web applies the same principles of the Web for digital content to the annotation of that content, proposing the re-use of unique, addressable URIs to refer to concepts just as URLs refer to content. In partic-

ular, the Linked Data initiative has helped mint unique identifiers for (almost) any concept in Web-wide accepted concept vocabularies, often associated to additional metadata about each concept that enables machines to further process annotations in support of more intelligent applications. Re-use of this existing Web based knowledge infrastructure is beneficial for Linked Television both due to its availability to any Web connected device as well as possibilities of shared understanding across distinct Linked Television applications which process the annotations.

Fig. 1. A example of a second screen application developed by LinkedTV for RBB news

In LinkedTV, we follow the Open Annotation Model which allows for flexible associations between annotation subjects (in this case, seed content fragments) and annotation objects (such as the concepts perceived as being represented in that fragment). Within LinkedTV's own annotation model (or ontology) which extends the Open Annotation Model, we can capture specific information about how annotations were created and how the seed content is fragmented, but particularly, we enable enrichments to be expressed as a specialized case of annotation. Since the target of an enrichment is itself a Web based resource, we

need nothing more than another URL to reference the content of an enrichment. As such, Linked Television as a concept is directly implementable using Web conforming specifications and data models.

3 Gaps in the broadcast TV integration

LinkedTV is designed as a new service for broadcast TV content, that is capable to provide added value enrichment of a TV program and provide accessibility to that enrichment synchronized to viewing. This is significant for the broadcasting industry which is coming under sustained competition from online content. We also identified the Web technology stack as being the most complete to enable media content annotation and to dynamically produce links between related media, considering experiments with HTML5 such as Mozilla Popcorn[5]. We observe that it is still more feasible to demonstrate media enrichments in the context of the Web than to apply it into the broadcast TV environment. Despite movement towards more Web technology support in SmartTVs and Set Top Boxes which are still the primary devices for classical TV broadcast consumption, media consumers seeking greater interactivity with content are more likely to consume that content over the Web than classical TV, despite initial "hybrids" solution such as second screen applications or hbbTV that promise the "best" of both worlds.

We identify the following gaps when developing the LinkedT concept:

- The media fragments URI specification identifies four dimensions to decompose a media resource: time, region, track and id. For LinkedTV, temporal fragments are by far the most significant dimension to identify the parts of the program where a topic is discussed or a particular object is depicted. Within the broadcast environment, an application needs a trustworthy reference to the correct temporal start/endpoint for the program in question, since an absolute timepoint creates redundancy for program repeats across channels and schedules, and a relative timepoint can be inconsistent if not measured in a consistent fashion (e.g. with or without commercial breaks). The media fragment URI specification provides the `clock` keyword for providing absolute time reference in a live stream. However, it lacks implementation from the web browsers.
- Applied to broadcast TV content, we create our own (abstract) URIs to identify a program, since the program description can be considered the same across distinct cases of program delivery (live broadcast, online catchup, web streamed). Each program instance can be attached in its metadata to different, network and time-dependant, addresses for accessing the digital representation of the program (locators). However, in broadcasting, the locators are often the only identification schemes communicated to devices and accessible via API to applications. On the other hand, any unique (abstract)

[5] `https://mozillalabs.com/en-US/Popcorn/`

identifiers used by content owners to identify their programs are not publicly visible, forcing Web applications which wish to additionally annotate and enrich broadcast media to mint their own URIs. The BBC has demonstrated an alternative approach with its Programme Ontology and use of Linked Data URIs to identify BBC programmes, which is worthy of wider promotion.

- Annotations, as defined by the Open Annotation Model, are relations between two or more resources (body and target), and their metadata, using an RDF graph. One of the major activities of the LinkedTV project is to use the Open Annotation Model to augment parts of TV content (as Media Fragment URIs) with annotations. These annotations can be other media resources, media fragments, keywords, concepts, LOD sources or arbitrary other online resources. Annotations can be either automatically detected or manually edited and assigned. Existing annotations give rise to further enrichments with additional annotations which can be domain or context specific. In a social TV context, tweets or comments can be related as annotations to media fragments. In LinkedTV, additional link sets are attached to fragments to enable synchronization between parts of a TV program and related online resources. Currently, LinkedTV generates such annotations within its own domain with content for which it has owners' permission (RBB in Germany and AVRO in The Netherlands) and can be re-used solely within its own workflow and applications. However, the supplemental information about the TV program captured by LinkedTV annotation goes far beyond what is today available to TV applications (largely EPG data). Yet, there is little effort currently made by the broadcast industry to either create or make available richer descriptions of their content despite the possibilities for new TV or second screen applications that would arise in the case of being able to access this metadata via appropriate APIs.
- The generation of annotated media fragments is currently mostly not possible in real-time, i.e. during a live broadcast event (except annotations via social streams), at least not for the indexing of the video track, since visual analysis techniques require a lot of processing. Using technologies such as Storm may in principle allow for a limited real-time processing of the audio track, or the live enrichment with related content during broadcast or VOD viewing. The main production scenarios, however, require that the video has either to be processed and prepared for annotation as a step of the production and publishing process, or else after the publishing for offering enhanced viewing options thereafter. This has to be taken into account within current media industry workflows.
- Applications able to access the annotations can use the information to support future Web and TV use case such as:
  - display of related content and advertising in overlay or multiscreen scenarios
  - create dynamic summaries of larger videos by assembling relevant fragments

However, this presupposes access to suitable functionality on device screens, such as playback of Media Fragments and overlays over video. Media Fragment playback is gaining support in the most recent versions of Web browsers but is not natively supported by Set Top Box media players or SmartTV platforms.

4 Proposals for standardization activity

We see a very strong case for pushing standardisation efforts for the Linked Television approach to seamlessly interlinking Web and TV, which requires support by CE manufacturers, content producers and broadcasters, to drive a new breed of innovative Web-TV applications and help SmartTVs and STBs to fulfil their promise of convergence of broadcast TV and Web. Among others, we propose the necessity to work on the following areas:

- general support of the Media Fragment URI 1.0 specification in media players in TVs and STBs
- consistent implementation of spatial and temporal markers for broadcast TV content in order to ensure a correct interpretation of Media Fragment URIs across devices and applications
- agreement on identification of TV content at an abstract level as opposed to one of its particular digital recording or stream. For the broadcast industry, such an agreement should ensure content owners maintain authority over the annotation of that content. Semantic Web and Linked Data best practices can provide a starting point for a decentralised approach.
- availability of richer TV annotations to applications running on TVs, STBs or connected second screens e.g. through the implementation of the appropriate `kind` role of the "Timed Annotation Track" defined in WebVTT or via another addressing scheme for finding (authoritative) annotations for some TV content
- introducing Annotated Media Fragments into productions standards and metadata exchange standards so that their creation begins as close to source as possible
- using Annotated Media Fragments in Multiscreen Environments, so that a second device could request and display a fragment of the content running on the first device
- synchronization between multiple devices, for example, following the proposal made within the W3C Second Screen Presentation Community Group

Acknowledgments

The research leading to this paper was partially supported by the European Union's 7th Framework Programme via the projects LinkedTV (GA 287911).