Deliverable 3.8   Design guideline document for concept-based presentations

Michiel Hildebrand, Lilia Perez Romero and Lynda Hardman

31/03/2015

Work Package 3: LinkedTV Interface and Presentation Engine

LinkedTV

Television Linked To The Web

Integrated Project (IP)

FP7-ICT-2011-7. Information and Communication Technologies

Grant Agreement Number 287911

Dissemination level[1]	*PU*
Contractual date of delivery	*31th March 2014*
Actual date of delivery	*31th March 2014*
Deliverable number	*D3.8*
Deliverable name	*Design guideline document for concept-based presentations*
File	*LinkedTV_D3.8.doc*
Nature	*Report*
Status & version	*Final*
Number of pages	*41*
WP contributing to the deliverable	*WP 3*
Task responsible	*CWI*
Other contributors	
Author(s)	*Michiel Hildebrand (CWI)* *Lilia Perez Romero (CWI)* *Lynda Hardman (CWI)*
Reviewer	*Lotte Belice Baltussen*
EC Project Officer	*Thomas Küpper*
Keywords	*Enriched video experiences, entity detection, companion applications, second screen, reusing web content*
Abstract (for dissemination)	*This document presents guidelines on how to setup enriched video experiences.* *We provide user-centric guidelines on the named entities that should be detected and selected to effectively enrich video news broadcasts. This is presented in the form of a user study. We selected 5 news videos and manually extracted the candidate entities from various sources, such as the transcript, visual content and related articles. An expert was asked to also provide interesting entities for the videos. The resulting 99*

candidate entities were presented to 50 participants via an online survey. The participants rated the level of interestingness of the entities and the usefulness of information from Wikipedia about these entities. Analysis of the results shows that users prefer entities of the type organization and person and have little interest for entities of the type location. They also indicate that subtitles are not enough as a source of interesting entities and that the amount of interesting entities can be improved by the combined use of subtitles with entities extracted from related articles or entities suggested by an expert. The expert suggestions showed to be more accurate than any other source of entities. Wikipedia seems to be a suitable source of additional information about the entities in the news, but should be complemented with additional sources.

We provide engineering guidelines on how to present, aggregate and process content for TV program companion applications. We describe the content processing pipeline that was developed in WP3 to feed the content for the Linked News and Linked Culture demonstrators. This shows how content from the Web can be re-purposed to enrich videos by extracting the core display content and presenting it in a uniform way to the user.

Table of Contents

1 Introduction

This document presents guidelines to create enriched video experiences. Within LinkedTV we developed three types of end user applications: An application for HbbTV compatible televisions and set-top boxes, where information is presented on top of a video broadcast and the user interacts with the content using the remote control; A second screen application that runs on a tablet device that is synced to the TV; and A TV program companion application that runs on a tablet and can cast video content to a TV.

With the single and second screen applications we focused on providing access to information about the entities mentioned or related to a TV program. For example, when watching a news item about the role of a minister in the construction of the Berlin airport the user can quickly look up background information about this minister. In particular for news programs the user studies provided evidence that users are interested in information about the persons, organizations and other entities related to the TV program. The evaluation of the prototype second screen application confirmed the added value background information about the named entities. *I saw a person in the Snowden story and I didn't know who he was, but I clicked the slide and found basic information to reassure my curiosity and help me understand* (LinkedTV Deliverable 3.5).

Essential to the success of these entity-driven applications is the selection of the entities that are presented to the user. Not all entities that occur in a news item are interesting for the user. For example, it is not necessary or desirable to show additional information about Berlin to Berliners every time the entity "Berlin" is mentioned in local news. To better understand which entities users are interested in when watching news videos we conducted a user study. The results of the study provide user-centered guidelines to apply entity detection and selection for enriched video experiences. We present the user study in Section 2.

In year 2 of the project we developed a prototype companion application for broadcasts of *international* news. Through user studies we identified the main interests of the participants: Explore how a news item is covered in other news sources (What does CNN say about this?); Access different opinions about the news item; Follow a news story throughout time (What has led up to this event?); Explore geo-localized information (What do people that live near the news event say about this?). In the prototype application we support these user interests. The application organizes similar type of information in so-called dimensions. For example, one dimension presents news articles about the same topic from online news sources, while another shows a timeline of past events.

In the year 2 evaluation the participants' reception of the application was very positive. *I like it, it is useful; It's so much better to have this tool than not to have this tool; The best is the simplicity. it's easy to use; it gives you easy access to different kinds of information with just one click* (LinkedTV deliverable D3.5). On average participants rated the usability of the

application according to the System Usability Scale (SUS)[2] with 83 out of 100. Participants in the study were generally enthusiastic to have different types of information related to the news concentrated in one application and considered that, in this sense, the interface would be time-saving with respect to their current equivalent practices (e.g. consulting several independent sources). The proposed exploration dimensions were described as very thorough in covering possible user information needs. The most valued dimensions were the timeline to follow a story throughout time, and the opportunity to read about the same news topic as it was presented by different media sources. Accessing geo-localized information through tweets was the functionality that participants liked the least.

In year 3 of the project we applied the companion application approach to the programs of the two content partners. We tailored the application to the cultural heritage program from AVROTROS „Tussen Kunst & Kitsch" in collaboration with the Netherlands Institute for Sound and Vision, and the regional News show RBB Aktuell in collaboration with RBB. The final companion applications were evaluated in the final user trials and are reported in LinkedTV deliverable D6.5.

The final user trials confirmed that the organization of related information into separate dimensions is well suited for companion applications. It allowed the participants to explore the topic from multiple perspectives, while at the same time it allowed users to select what best matched their interests and skip the dimensions they were not interested in. The studies also showed that the participants like that the application provides an aggregation of information from multiple sources within a single application. To realize this aggregated presentation of Web content into a single application we developed various content processing tools that we collectively refer to as the content proxy. The input for these tools are the locators of the content generated by the LinkedTV system, typically URLs, and the output is a uniform representation of the content found at the URL that can be used by the companion application. In Section 3 we explain the components of the content proxy and therewith provide engineering guidelines on how to re-use Web content in enriched video applications.

Finally we wrap up the document with an overview of the applications and services developed within WP3 in Section 4.

---

[2]   http://en.wikipedia.org/wiki/System_usability_scale

2 User-centered guidelines for finding relevant entities to use in web enriched news videos

Entities play an important role in the LinkedTV applications. The single and second screen applications focused on providing access to information about entities to help users answer questions such as: Who is this person? Where is this organization located? To effectively support users of video applications with information about named entities we need systems that not only detect named entities, but also determine which of the detected entities are *interesting* to the user to receive more information about. For example most TV news viewers would not be interested in reading the Wikipedia page of Barack Obama after watching an item on the Obama Care plan, or the US involvement in Middle Eastern conflicts.

Brzeski et al. [2] distinguished *user-centric* entity detection systems from *machine-centric* systems and introduced *interestingness* as a quality dimension that is essential for user-centric systems. An entity is interesting when it is "useful or compelling enough to tear the user away from the main thread of the document". To date, most research regarding named entity detection from video content focuses on the machine-centric approach. Little attention has been given to user preferences regarding which entities they consider worthy of acting as anchors to additional information (degree of *interestingness*). For example, most entity detection systems for video content try to detect the entities in the closed captions, the visual content (faces) or textual labels, while it is not clear if these sources actually contain the entities that interest the users.

The intention of this research is to better understand the characteristics of the entities that users consider interesting in the context of news videos. The results of this study have implications for the design of *user-centric* entity detection and selection strategies in the field of web enriched news videos. Specifically, we study from which sources entities are best extracted, (e.g. closed captions, visual content), which types of entities (persons, organizations and locations) interest users and how interesting entities relate to the topic of a news item. In addition, we study if Wikipedia provides the information about the entities that users consider useful.

We chose to focus on news videos, as the informative nature of this genre is well suited for the presentation of additional content. Furthermore, information about the news appears daily in diverse media formats, providing richness and variety to the hypermedia experience. The layout in which the news is presented in newscasts is fragmented in many, often unrelated stories, therefore, presenting additional information would seem natural in a news context and would not interrupt the story line as it would in other genres [3].

For the study we created a dataset of candidate entities for 5 news videos by manually identifying them from different sources and annotating them with their type and their relation to the topic of the video. We also asked an expert to provide the interesting entities for each video. With an online survey we collected feedback on the interestingness of these named

entities from 50 participants. Using the dataset and the participants' feedback we perform an analysis of the characteristic of the interesting entities.

In the next section, we first present related work before we describe our research questions. In the fourth section we explain our method and describe the user survey that we conducted. This is followed by the enumeration of the results that we analyze in the discussion. We conclude with a reflection on the lessons learned and future work.

2.1 Related work

Brzeski et al. [2] present and evaluate an algorithm by showcasing it in the "Contextual shortcuts" system (a large scale user-centric entity detection platform). They demonstrate that using context-surrounding entities can greatly improve user-centric systems performance in relation to user preferences. In this work, the extent to which the algorithm matches users' preferences is considered the main measure of success. The authors distinguish between user-centric and machine-centric entity detection systems, and argue that for the user-centric systems the simple precision/recall scores are not suitable because they don't reflect the quality of detected entities. As a solution they propose to measure the quality of detected entities within a user-centric entity detection system using three core dimensions: the accuracy, the interestingness and the relevance of the entities it presents to the users. Their motivations coincide with ours in that they identify how entity detection targeted to users should be different from entity detection targeted to machine use. They place users' preferences regarding entities in the center of their concerns. In this study we adopt their definition of "interestingness", and we investigate some aspects of the entities considered interesting. In our study accuracy and relevance are not measured because we used manual detection of entities.

Web enriched news is a subject that has been researched as a use case in various studies concerned with hypermedia and information retrieval [4][5]. Named entity extraction is one of the techniques used to automate news annotation and retrieval [6][7][8].

Often integration of the newscasts with the Web occurs through linking fragments of texts in the newscasts captions with Wikipedia pages.

Milne et al. use Wikipedia human-made links to train an algorithm than can identify significant terms within unstructured text, and enrich it with links to the appropriate Wikipedia articles [1]. One of their use cases is to enrich news broadcasts. The authors use Mechanical Turk human evaluators to rate the links obtained by the algorithm in terms of whether they are the right Wikipedia articles about a word or term found in the newswire. They take into account whether the evaluators consider the Wikipedia article to be relevant and helpful. Only relevant and helpful links are considered correct.

Odijk et al. [9]describe and evaluate a solution to the task of real time semantic linking streaming text and in concrete for the use case of a news talk show. Their solution includes a learning to re-rank approach that significantly improves over a strong baseline in terms of effectiveness.

Henzinger et al. [10] investigated the possibility of showing related news articles to users while they watch news on TV. Their approach is the timed based keyword extraction and works by sending fragments from an ongoing stream of closed captions to a news search engine.

All three relate to our work in their use of Wikipedia and related articles to enrich the news broadcast. All three examples use human annotators to create a ground truth and distinguish relevant or interesting links and their correctly matching anchors. However their motivation in all cases diverges from ours because their main goal is to evaluate and present a series of algorithms and semantic linking techniques rather than to investigate user preferences. Their evaluations address the system's accuracy in creating links. We want to explore some qualitative characteristics of the anchors (entities) that produce the links, especially the ones considered interesting by the users.

In the field of human computer interaction and user experience, multiple works have described the design and implementation of systems for the interactive and enriched consumption of broadcast news. Ardissono et al. [11] investigated personalization in the context of interactive news broadcasts through tracking and inferring user content interests and media preferences.

MyNewsMyWay is a system for nonlinear and interactive news broadcasts with an emphasis on the user needs [12]. With the intention of recreating the television experience that users are most acquainted with, MyNewsMyWay uses the remote control as input device and displays the system interface on a TV screen. A similar remote control, single screen set up was produced and evaluated by [13] to enable users to access opinions over stories headlines, skip back and forth over stories and select stories from a playlist. Also through a remote control [14] gives users access to summaries about the video news and to a list of stories prioritized using metadata and the use profile.

In our previous work [15] we explore the consumption of enriched newscasts through the use of a second screen device.

The aforementioned HCI works are useful in providing a context for the current study, in the sense that they demonstrate possible applications of semantic linking through entities. They coincide with our work's motivations in that they focus in exploring users' needs. However, they explore aspects of the interaction and the user experience. To the best of our knowledge none of them focus in exploring user's preferences regarding entities.

2.2 Research questions

1. How does user interest in entities differ by entity type?
2. Are user interests covered by entities extracted from the subtitles/transcripts alone?
3. Can experts help in selecting/filtering interesting entities?
4. Are users interested in entities that are the main topic of a news item?
5. Is Wikipedia sufficient as the source for information about named entities?

Question 1 looks at entity types in relation to interestingness. Do users consider entities of a certain type more interesting than others or is there no relation between entity type and "interestingness"?

Question 2 looks at the entities that are rated as interesting and whether all of them can be found among the named entities detected within the speech transcription or subtitles which are the most common source of entities and key words used in most research [9][10][13][14][18]. It also investigates whether there are other potentially useful sources for providing interesting entities. In particular we look at the following sources of named entities: video image [16]; text contained in the video (e.g. nametags or banners) [17]; documents related to the subject of the video [18][19]; and entities suggested by an expert [20].

Question 3 deals with whether human annotators or editors can help in the labor of pointing out interesting entities. It relates to works like [20] who suggest that despite the efficiency of a semantic linking algorithm, a human filter is needed before delivering contents to the user. We look at entities suggested by an expert, how users evaluate them, and whether they add to what can be automatically extracted.

Question 4 investigates whether entities contained in the video title, which usually describes the news item, or entities that respond to the fundamental questions *who, what* and *where* tend to be considered interesting by the users or not, or whether there is no correspondence between both.

Question 5 examines whether wikifying (linking anchors to Wikipedia articles), a resource used by many researchers like [1][9], is supported by users interest. In other words, we look at whether newscasts viewers would like to receive additional information from Wikipedia. We also look at whether there is a relation between the entity type and their preferences regarding Wikipedia as a source of information.

2.3 Experiment

We selected five short videos from BBC One Minute World News. The selection covered a wide range of subjects specifically: politics, armed conflicts, environmental events, legal disputes, and social news. Since the user study was going to include international participants, another criterion for the selection was to choose news that had caught the attention of international audiences.

We chose to focus only on entities of the types person, organization and location because they can directly translate or answer three questions: Who is involved? What happened? Where did it take place? These questions are a subset taken from a well-known formula in journalism known as the 5Ws: *who*, *what*, *when*, *where* and *why*, which emphasizes the fundamental dimensions that an informative journalistic text should report on [21]. Regarding the questions *when* and *why*, they were discarded because in order to acquire meaningfulness they need to be modeled not only by single entities but also by more complex relations between them that are out of the scope of the current paper.

Subtitles of the videos were not available; therefore, a member of the team manually transcribed the speech in the videos. After obtaining the transcriptions, the following steps were performed in order to obtain an unbiased set of candidate entities.

2.3.1 Transcripts

All entities of the type person, organization and location were manually extracted from each one of the video transcripts and added to the unfiltered list of entities (candidate set).

2.3.2 Image in the video

The video image was visually analyzed by a researcher and every time a recognizable person, organization or location was portrayed this was also added as an entity to the candidate set.

2.3.3 Text in the video image

The video was analyzed for text appearing in the image. Whenever text appeared in the video image, for example in the form of nametag overlays, the named entities appearing in such tags were added to the candidate set.

In order to complement the candidate set with entities that might be interesting for the user, but are not necessarily found in the videos, we used the following two strategies:

2.3.4 Suggestions of an expert

We collaborated with a journalist with more than 6 years of experience as a writer/editor for important American newspapers and websites. We configured an online survey to retrieve the expert's feedback. In the survey we explained what named entities are and which types of named entities we needed. After this introduction we presented the videos to the expert. Following each one we asked him to list the named entities that, according to his criteria, would better serve the objective of showing interesting additional information to the users. The expert didn't have access to the candidate set, and was completely free to suggest any named entity he wanted.

2.3.5 Related articles

We created a Google custom search for three news sources: The Guardian, New York Times, and Al Jazeera online (English). Using this search engine we looked for articles related to the video. We performed the search using the main terms in the videos' title, for example, for "Fugitive Edward Snowden applies for asylum in Russia" we searched for "Edward"+"Snowden"+"asylum"+"Russia". We limited the results to +- 3 days from the day when the video was published. We chose one document from each source, the one closest in topic and time to the video. We then extracted all named entities of the types person, organization and location from the resulting documents. In order to keep the number of entities within a reasonable amount for inclusion in a survey, we kept only the named entities

that appeared in at least 2 related articles and dropped all the ones that only appeared in one. The selected entities were added to the candidate set.

2.4 Refining the candidate set

We refined the candidate set comprised of all found entities by eliminating all named entity duplicates and standardizing names. For example, when we had "Barack Obama" as an entity and "Obama" as another entity we eliminated the shorter one and left the complete name. The end result is a set of 99 entities for the five videos. Table 1 shows the distribution of the entities among types and videos. The number of entities varies per video from 13 for the „Fukushima" video and 28 for the „Snowden" video. The majority of the entities are of type person (46). We obtained 20 organizations and 33 locations.

Video's Title (News Item)	Person	Organization	Location	Total
Fugitive Edward Snowden applies for asylum in Russia	11	7	10	28
Egypt's Morsi Vows to Stay in Power	7	5	5	17
Fukushima leak causes Japan concern	4	5	4	13
Rallies in US after Zimmerman Verdict	9	2	8	19
Royal Baby Prince Named George	15	1	6	22
	46	20	33	99

Table 1: Distribution of manually identified candidate entities per video and type.

Online Survey

We created an online survey with the aim of gathering information about the degree of *interestingness* of the entities in the candidate set.

Fifty international subjects participated in this online study. They responded an online call distributed via email and social networks. Their age range was between 25 and 54 years with an average age of 30.3 (standard deviation 7.3 years). 18 participants were female and 32 were male. Most of the participants were highly educated and 48 of them had either a university bachelor degree or a postgraduate degree. The main requisite for participation was that they were interested in the news and followed the news regularly, preferably through means that include newscasts.

During the survey participants were asked to choose at least 3 out of 5 videos according to their preferences. Then they were shown each one of the videos. After each video, they were

asked to rate whether they would be interested in receiving more information about the named entities in the context of the news video, on a second screen or similar application. All the related entities from the candidate set were shown in a list with ratio buttons arranged in a similar way to a three-point Likert-scale. The possible answers were "Yes" "Maybe" and "No". For each entity they selected (maybe or yes) they also had to indicate if they would like to get the information about this entity from Wikipedia. The possible answers were "Yes" and "No".

2.5 RESULTS

With the online survey we collected the degree of interest in the selected entities from 50 participants. As the participants could select three videos that interested them the number of respondents varies per video. From the 50 participants 44 selected three videos and six participants selected all five videos. Table 2 shows the number of respondents and the total number of judgments per video. In total we collected 3.218 judgments over 99 entities. The „snowden" and „fukushima" videos were selected most. The video about the „Royal baby Prince" was selected 14 times.

Video's Title (News Item)	Entities	Respondents	Judgments
Fugitive Edward Snowden applies for asylum in Russia	28	45	1260
Egypt's Morsi Vows to Stay in Power	17	34	578
Fukushima leak causes Japan concern	13	42	546
Rallies in US after Zimmerman Verdict	19	27	513
Royal Baby Prince Named George	22	14	308
	99		3205

Table 2: Number of respondents and the total number of judgments per video.

Table 3 shows how the judgments are divided over the categories *yes, maybe and no.* A negative judgment (*no)* was given in a bit more than half of the cases (52%). In 26% of the cases the judgment is *maybe* and in 22% of it is *yes.* While the number of respondents varied per video the distribution over the answers is comparable for all five videos. The low number of 15 respondents for the „Royal Baby" video did not effect this distribution. Therefore we decided we could use the judgments for all five videos in the result analysis.

Video's Title (News Item)	Yes	Maybe	No	Total
Fugitive Edward Snowden applies for asylum in Russia	231 (18%)	314 (25%)	715 (57%)	1260
Egypt's Morsi Vows to Stay in Power	157 (27%)	173 (30%)	248 (43%)	578
Fukushima leak causes Japan concern	132 (24%)	144 (26%)	270 (49%)	546
Rallies in US after Zimmerman Verdict	95 (19%)	119 (23%)	299 (58%)	513
Royal Baby Prince Named George	78 (25%)	86 (28%)	144 (47%)	308
	693 (22%)	836 (26%)	1676 (52%)	3205

Table 3: Distribution of the judgments over the answers yes, maybe and no.

The positive judgments are divided over the total set of entities. The chart in Figure 1 plots the distribution of the judgments for each entity. There are clear tendencies in which entities are considered interesting by most participants, such as Edward Snowden in the „Snowden" video and Muslim Brother in the „Egypt" video. To determine the central tendency for each entity we computed the mode. The mode is a statistical measurement for the value that occurs most often. For example, the entity Edward Snowden was judged *yes* **30** times, 10 times *maybe* and 5 times *no.* The mode is yes. In the chart (Figure 1) the green bar for Edward Snowden (second row from the bottom) is predominantly green. There are 19 entities out of the 99 for which the mode is *yes* and for 8 entities the mode is *maybe.* For the other 72 entities the mode is *no.*

The 27 entities for which the mode was *yes* or *maybe* are interesting for most users and we refer to this set as the *interesting* or *selected* entities. For each entity we also computed the average score. We mapped the judgments to numerical values, 1 for *yes, 0* for maybe and -1 for no, and computed the average score per entity over all participant judgments. We use this average measure to complete the analysis of the results based on the set of selected entities.
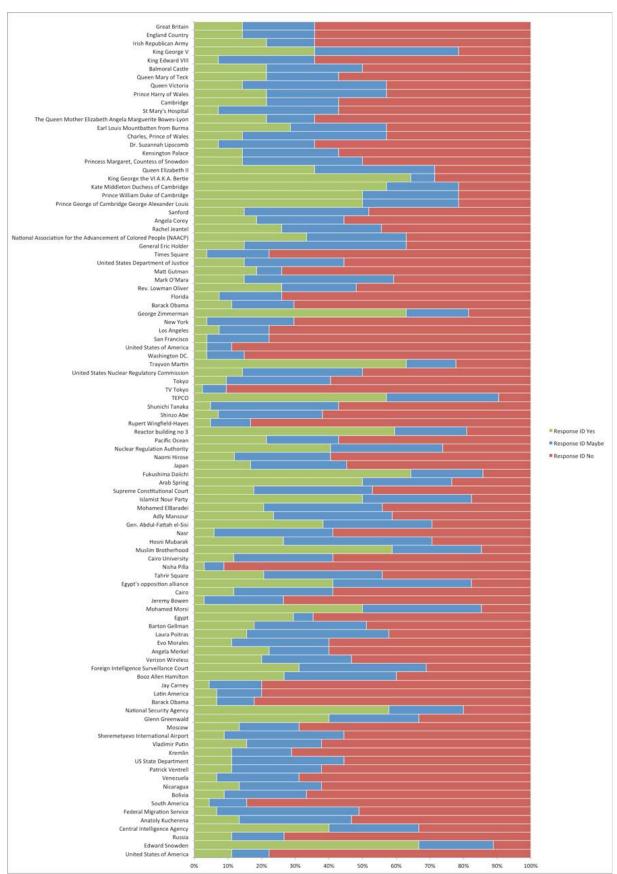
Figure 1: Aggregated user judgements for all 99 entities.

2.5.1 User interest by entity type

From the set of selected entities the general tendency of the participants shows there is very little interest in entities of type location and proportionally the highest interest is in organizations. Table 4 shows the distribution of the selected entities over the entity types person, organization and location. Only 1 of the 27 selected entities is a location. Proportionally organizations were most often considered interesting. In half the cases an organization was considered somewhat interesting (*yes or maybe)* and persons in 35% of the cases.

Type	Total	Selected (*yes*)	Selected (*yes* or *maybe*)
Person	46	11 (24%)	16 (35%)
Organization	20	8 (40%)	10 (50%)
Location	33	1 (3%)	1 (3%)
	99	20 (20%)	27 (27%)

Table 4: Proportion of types for selected entities by mode yes and by mode yes or maybe.

Type	Total	Yes	Maybe	No	Avg.
Person	1354	300 (22%)	368 (27%)	**686 (51%)**	-0.2
Organization	763	**252 (33%)**	228 (30%)	283 (37%)	**0.0**
Location	1088	141 (13%)	240 (22%)	**707 (65%)**	-0.5
	3205	693 (22%)	836 (26%)	1676 (52%)	-0.3

Table 5: Total number of judgments per entity type, the average score and the number of judgments per value.

When inspecting the locations in more detail the selected location is the entity *Reactor building no3* in the „Fukushima" video. This is not an administrative location such as a city or country, but a specific point of interest. The total set of 33 locations contained 4 other specific points: Kremlin, Sheremetyevo International Airport, Cairo University and Tahir square. All other locations were either a country, city. We did not find a correlation between the specific points and the users' interest compared to the other places or cities. All non-selected locations have a negative average score close to -0.5.

The total number of judgments per entity type confirms the tendency reflected by the set of selected entities. Table 5 shows the number of judgments per entity type over the values yes, maybe and no. The lack of interest for locations is also shown here with 65% of the judgments being negative (no). The total average score for locations is -0.5 versus -0.2 for persons and 0.0 for organizations. For persons more than half of the judgments are also negative (51%). Only for organizations the majority of the judgments is positive (maybe or yes).

2.5.2 Coverage of user interest by entity source

The participants consider not all entities extracted from the subtitles interesting, and from the total set of selected entities less than half can be extracted from the subtitles. Table 6 shows for each of the five sources the total number of entities, the entities from that source in the set of selected entities and the coverage the source provides over the total set of 27 selected entities. From the 40 entities that were extracted from the subtitles 13 (33%) are in the set of selected entities (*yes* or *maybe*). From the total of 27 selected entities these 13 entities from the subtitles provide coverage of 48%. The visual content is comparable to the subtitles in terms of selected entities as well as coverage. Most of the selected entities are contained in the related articles (78%).

Source	Total	Selected (yes or maybe)	Coverage	Not in subtitles
Subtitles	40	13 (33%)	48%	-
Text overlays	9	2 (22%)	7%	1
Visual content	35	12 (34%)	44%	1
Expert	37	18 (**49%**)	67%	9
Related articles	69	21 (30%)	**78%**	10
	99	27 (27%)		

Table 6: Selected entities per source

The added value of a source in terms of coverage is expressed in the final column of Table 6. The text overlays and visual content both add only 1 selected entity in addition to the subtitles. The expert contributed 9 additionally selected entities and the related articles contributed 10 entities to the subtitles. Figure 2 illustrates the overlap among the 5 sources. Most of the selected entities were found in more than one source. The text overlays and the visual content both contained the same entity Mark O'Mara in the Travyon Martin video. The related articles contributed five entities that were not contained in any other source. The expert contributed three unique entities and the subtitles contained one entity that did not appear in any of the other sources (Egypt's opposition alliance).
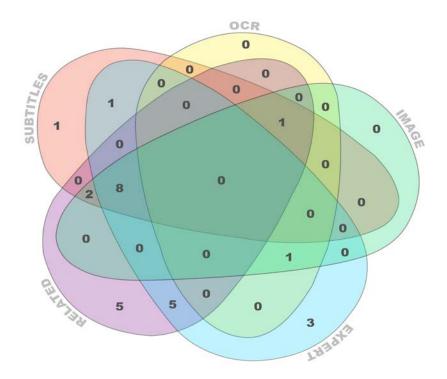
Figure 2: Distribution and overlap of the selected entities over the five sources

2.5.3 Expert judgments

The entities from the expert provided the highest precision compared the set of selected entities (49%). The expert provided coverage of 67% missing 9 out of the 27 selected entities. Four of these entities were contained in the subtitles and the other 5 in the related articles. The expert provided 3 entities that were considered interesting by the majority of the participants and were not contained in any of the other sources.

Half the entities (51%) provided the experts were **not** considered interesting by the majority of the users. We analyzed these entities individually. We discovered that the expert selected locations for each of the five videos and the participants selected none of these. Our hypothesis is that the expert selected these locations by their relevance to the news item. For locations relevance to the topic is not a criterion for interestingness. We also observed that the expert selected several persons and organizations that were not found in any of the other sources. Here our hypothesis is that the expert knew these entities from his domain knowledge, but to the participants these entities were not known and too far away from the topic of the news item. Entities added by the expert that were also found in another source, we speculate that these are thus closer to the topic, were more often selected by the participants.

The experts were not given the set of candidate entities, but instead they were asked to provide the entities themselves. In an editorial setting where these entities are given we expect the expert to perform better on coverage.

2.5.4 Interest in the main topic of the news item

Participants are interested in information about persons and organizations that are the main topic of the news item. We considered an entity to form part of the main topic when it answers one of the fundamental questions: Who is the news about? What is the news about? Where did the news happen? Our candidate set contained in total 11 entities that we manually identified as being part of the main topic of the news item. Table 7 shows per video the entities that were identified as being part of the main topic. Six out of the ten entities that formed part of the main topic were considered interesting; all of these were of type person (5) or organization (1). The entities that were not selected as interesting were all of type location (4).

From the 10 entities selected as the main topic, 9 of them were also found in the title of the news item provided with the video. Only the entity Trayvon Martin in the „Zimmerman" video was not included in the title.

While being part of the main topic (the title) is a good indicator for interestingness this only applies for 6 out of the 27 selected entities. In other words 21 of the selected entities are not part of the main topic.

Video	Entity	Type	Mode
Fugitive Edward Snowden applies for asylum in Russia	Edward Snowden	Person	Yes
	Russia	Location	No
Egypt's Morsi Vows to Stay in Power	Mohamed Morsi	Person	Yes
	Egypt	Location	No
Fukushima leak causes Japan concern	Fukushima Daiichi	Organization	Yes
	Japan	Location	No
Rallies in US after Zimmerman Verdict	Trayvon Martin	Person	Yes
	George Zimmerman	Person	Yes
	United States of America	Location	No
Royal Baby Prince Named George	Prince George of Cambridge George Alexander Louis	Person	Yes

Table 7: Entities in the main topic of the news item

2.5.5 Wikipedia as an information source

When users are interested in information about an entity, Wikipedia is useful in less than half of the cases. Table 8 shows the total number of entities per type that users selected. We consider the total number of users that judged an entity *yes* as well the combination of judgments *yes* or *maybe.* The column labeled **Total** contains the number of positive

judgments and the column labeled **Wikipedia** shows the number of times the user indicated Wikipedia as a useful source for the selected entity. Considering the combination of *yes* or *maybe* in 43% of the cases that a participants judged an entity interesting he/she indicated that information from Wikipedia would be useful. For organizations this is 49% and for locations 39%. Considering only the *yes* judgments these numbers are higher. In 60% of the cases the participants indicated that information from Wikipedia is useful. The relative numbers per entity type are also higher. Entities of type person and organization are above 60%.

We conclude that Wikipedia is a useful information source when a user finds an entity interesting. It is, however, not the case that information from Wikipedia is useful for each interesting entity and for each user. For only 20 of the 27 selected entities (*yes* or *maybe*) the majority of the participants stated Wikipedia to be useful. The average number of participants that find Wikipedia useful per selected entity is low with 53%. Therefore, we also conclude that information about entities should not only come from Wikipedia. It remains future work to investigate in which cases (which entities and which users) Wikipedia is useful and which sources can complement Wikipedia with respect to information about entities.

Type	**Judgments** (yes)		**Judgments** (yes or maybe)	
	Total	Wikipedia	Total	Wikipedia
Person	300	183 (61%)	668	268 (40%)
Organization	252	162 (64%)	480	237 (49%)
Location	141	75 (53%)	381	147 (39%)
	693	420 (60%)	1529	652 (43%)

Table 8: Usefulness of Wikipedia as an information source. The total number of entities per type judged as interesting and the number of times the user indicated Wikipedia is useful for that entity.

2.6 DISCUSSION

The initial candidate set of entities that we obtained by manually extracting named entities had a total of 99 entities. These were presented to participants in our survey who rated neutrally or positively (with a majority of "maybe" or "yes" only 27 or them). This indicates the need for filtering the named entities that can be extracted taking into account users' preferences. Participants in the study stated that they would not like to receive additional information about most of the entities in the original candidate set, therefore, showing

additional information about all entities would result in an overload of information in most cases considered irrelevant by users.

RQ1. In terms of entity type, our participants seemed to prefer to receive information about organizations (they found 50% of the organizations relevant), to persons (35%). Locations were by far the entities considered the least interesting (3%).

RQ2. Although subtitles or speech transcriptions are the most frequently used source of entities for video enrichments, in our study they did not seem to be the most thorough. Using only entities detected in the subtitles we would have just found 48% of the interesting entities. In contrast 78 % of the entities would have been found using only related articles and 67% using only the expert suggestions. From the entities found in the subtitles a third were found useful by the majority of the users. .

The results of our experiment suggest that using subtitles in conjunction with other sources may greatly improve the performance of a system in terms of how thorough it is in retrieving relevant entities. In particular this supports the strategy proposed by [18][19] which consists in expanding the named entity extraction from the subtitles to related documents. In the case of our study using newspaper articles from 3 different newspapers about the same news item, improved the thoroughness from 48% (subtitles alone) to 85% (subtitles combined with related documents).

RQ3. Although manually enriching news videos may be too time consuming, our findings seem to suggest that a hybrid machine-human solution may be worth exploring. An automated system can suggest entities and enrichments to a human editor that can play the role of filtering them according to user's preferences and suggesting new ones. In our study, when combining the entities found in the subtitles with the suggestions of the expert, 81% of the interesting entities could be found. Likewise, the suggestions by the expert were the most accurate or precise because 49% of the suggested entities ended forming part of the group selected as interesting.

RQ4. In our study, all the persons and organizations that were part of the main topic were considered interesting by the majority of the participants. On the contrary entities of the type location were not considered interesting regardless of whether they formed part of the main topic of the news.

RQ5. According to our study, Wikipedia seems to be a suitable source for providing additional information about entities in the newscasts. When a participant rated an entity as interesting he or she was often also interested in information from Wikipedia. Considering ratings from *yes* or *maybe* Wikipedia was considered useful in 43% of the cases and when considering only *yes* rating in 60%. Wikipedia should be complemented with additional data sources, as not all participants consider. For 20 out of the 27 selected entities the majority of the participants indicated Wikipedia as useful.

2.7 Future Work

Our study was relatively small, including only 5 news videos and 50 participants. Further studies with a larger amount of both would be desirable. In future studies, more entity types should be included such as field-specific terminology that may complement the information about the news. One thing that may be interesting to take into account for future work is how users' preferences may vary depending on the category of the news being portrayed. For example, users' preferences regarding news about entertainment and spectacle may raise different doubts and needs than news about natural disasters.

More complex relations between entities can answer the fundamental journalistic questions that were left out of this study: *Why did this happen? How did this happen?* They could also contribute to give meaningfulness to the question *When did it happen?* Performing a user study where complex relations between entities are included would be interesting as future work.

3 Engineering guidelines for reusing Web content in enriched video applications

The LinkedTV companion application is a tablet application that can be tailored to a specific TV program. We tailored the application to the Dutch program *Tussen Kunst en Kitsch (TKK),* courtesy AVROTROS and to the *RBB aktuell* news show. These applications were used in the final user trials. Details of the user interface and configuration of the dimensions are available in LinkedTV deliverable 3.7. We recap the main parts of the application and the presentation of the enrichment content before we describe how the content is generated for them.

3.1 Recap of companion application design

The application contains three parts. First the user selects an episode from the overview page (Figure 3). Second the user watches the video in the player (Figure 4). The user can watch the video on the tablet or cast it to a TV. The user can skip to another chapter by selecting one from the chapter list. Third, the user opens the *explore* mode of the application when she wants to know more about the content discussed in the chapter (Figure 5). The explore mode can be opened while the chapter is playing or when it is finished. If the user is watching the video on the tablet the video is paused while the explore mode is opened. If the user has chosen to cast the video to the TV the tablet operates as a second screen and the explore mode is accessible simultaneously with the video playing on the TV. The explore mode gives access to background information and related content. Similar information is organized in so called dimensions. In the application these are shown on the left side of the interface. Each dimension is represented by a blue header. Selecting the header reveals the items in this dimension. In Figure 5 the background (Achtergrond in Dutch) dimension is opened and an article is selected. The content of the article is shown as the main content on the right side of interface.

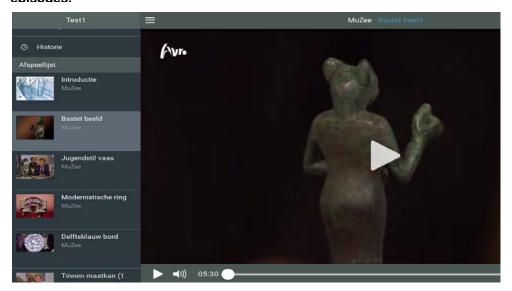**Figure 4: Screenshot of the companion application for Tussen Kunst & Kitsch showing the episodes.**

Figure 3: Video player and on the left side the chapter navigation.

Figure 5: Explore mode showing the dimension on the left side and the content of the select background article on the right side

3.2 Presentation of enrichments

The explore mode of the application gives access to various types of information coming from different sources on the Web. Figure 7, Figure 8 and Figure 9 show content that was selected by an editor for one of the chapters from a Tussen Kunst & Kitsch episode. Respectively, a page from DBPedia over the concept *Baset*, a Bastet artwork from Europeana, and Web page about the use of bastet statues in ancient Egypt. Each source has its own layout, design, branding, navigation and advertisements. A straightforward way to present this information would be to directly include the original source in the application. This, however, makes it difficult for the user to quickly explore content from different sources, as she needs to adapt to the layout and design of each source. Furthermore it is more complicated to focus on the actual content, as information not related to the content such as navigational elements and advertisement are distracting. To simplify the user's exploration we need to present the information from different sources in a uniform way, similar as the popular aggregation App Flipboard[3.]

To achieve this we developed several tools and services to fetch and extract content from different types of Web resources. To feed the extracted information into the application and present it in a uniform way we defined a data model for the application. The application uses this data model to select the different pieces of information and present them in the right place in the interface.

In Figure 9, Figure 10, Figure 11 show the same content as selected by the others, but now after processing and embedded in the application. Figure 9 shows the information about the entity *Bastet.* Figure 10 shows the Related artwork and Figure 11 the background article. In addition, the application gives access to related chapters from other episodes of Tussen Kunst & Kitsch. These chapters are presented in a similar way as the other types of content as shown in Figure 12.

The presentation of enrichments in the application is similar for all kinds of sources, but it allows variation to tailor it for specific types of content. At the top it shows the title, the source of the content, such as the name of the museum in case of the background article. On the right side in the header there is a link to the original source. In the main area below the header the content presentation depends on the type of the content. Web pages, such as found in the background dimension are presented by a big image (or video if available) and the body text. Entities, artworks and related chapters are presented by an image and next to it a short description and metadata properties.

We describe the data model and the content extraction services in turn.

---

[3]   http://flipboard.com

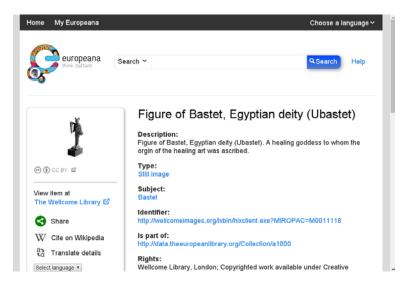**Figure 6: Webpage from DBpedia showing content about the entity *Bastet*.**

Figure 7: Webpage from Europeana showing content about the Bastet artwork.

Figure 8: Webpage with the article about the use of goddess statues in ancient Egypt.
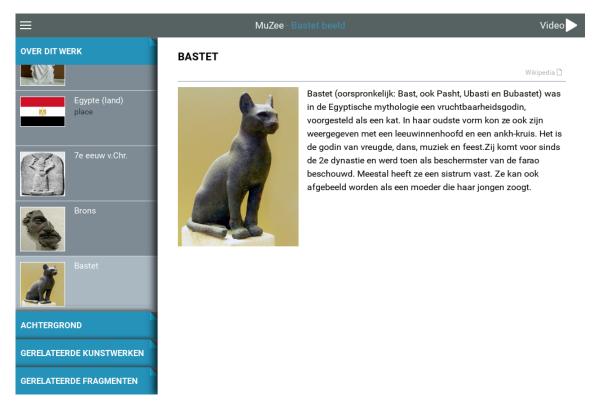
Figure 9: Enrichment from the about dimension, showing information about the concept *Bastet*
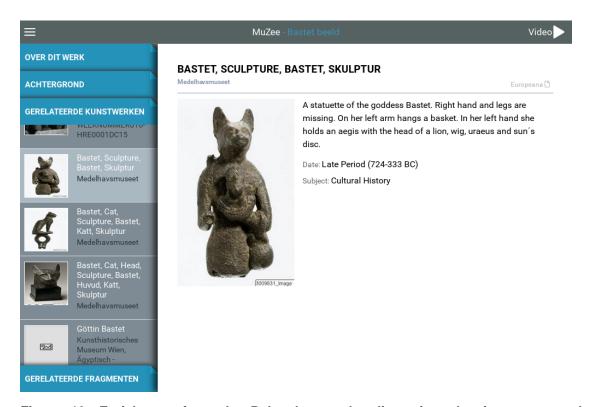

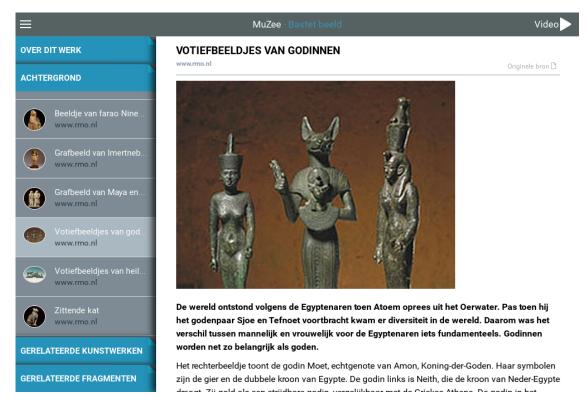**Figure 10: Enrichment from the Related artworks dimension showing an artwork from Europeana.**

Figure 11: Enrichment from the background dimension showing an article about the role of godesses statutues in ancient Egypt.

Figure 12: Enrichment from the related fragment dimension showing a TKK fragment about an Egyptian mask.

3.3 Application data model

The data model of the application contains four levels: information about the program, the chapters in this program, the dimensions in each chapter, and the items in each dimension. For a program the application currently reads in a JSON file that contains the information at all four levels. The LinkedTV news and culture demo applications use static JSON files. The application can be extended to read the data dynamically from a web service.

At the top level the model contains information about the program (Figure 13). This information is used in the overview of the available episodes and in the video player of the application. It requires an identifier (*id)* of the program, a *title* and the source of the video stream (*src*), and a *poster* image representing the video. The list of chapters shown in the video player is taken from the *chapters* array.

```
{
    src: "http://player6.noterik.com/stream17/domain/linkedtv/user/avro/video/104/rawvideo/2/raw.mp4",
    curated: true,
    title: "MuZee",
    poster: "http://images1.noterik.com/domain/linkedtv/user/avro/video/104/shots/1/h/0/m/0/sec10.jpg",
  - chapters: [
      + {…},
      + {…},
      + {…},
      + {…},
      + {…},
      + {…}
    ],
    shots: "http://images1.noterik.com/domain/linkedtv/user/avro/video/104/shots/1/",
    id: "8a8187f2-3fc8-cb54-0140-7dd2d0650005"
}
```

Figure 13: Datamodel with information at the program level

A chapter (Figure 14) is represented by the start (*startTime)* and end time (*endTime*). In addition, an identifier (*id*) and a *title* are required. If an *image* is provided this is used to represent the chapter in the application. In case this is not available an image is taken from the shots service provided in the program model. The application uses the start time of the chapter to fetch an image from this service. The dimensions shown in the explore mode of the application are taken from the *dimensions* array of each chapter.

```
- {
    + dimensions: [...],
      title: "Bastet beeld",
      image: "http://images1.noterik.com/domain/linkedtv/user/avro/video/104/shots/1/h/0/m/8/sec11.jpg",
      startTime: 330003,
      duration: 241000,
      endTime: 571003,
      id: "8a8187f2-3fc8-cb54-0140-7dd2d0650005#t=330.003,571.003"
  },
```

Figure 14: Data model with information at the chapter level

A dimension (Figure 15) is represented by an identifier (*id*), a *title* and a *type*. The title is used in the header of the dimension list. The items in a dimension are provided in the *items* array. The presentation of an item in the dimension list and in the content pane is determined by the type of the dimension. The application currently supports *entit*y, *article*, *artwork, youtub*e and *chapte*r. The application optimizes the presentation of the content for these types. The application can be extended to support other types of content by providing HTML templates for them.

```
- dimensions: [
    - {
        + items: […],
          type: "entity",
          id: "maintopic",
          title: "Over dit werk"
      },
    - {
        + items: […],
          type: "article",
          id: "tve_1",
          title: "Achtergrond"
      },
    + {…},
```

Figure 15: Data model with information about a dimension

An item (Figure 16 and Figure 17) in a dimension also contains the obligatory *url*, a *title* and a (thumbnail) *image*. Optionally the *source* and an *author* can be included. The data model for an item depends on the content type. Figure 16 shows the object-based model that is used for the content types *entity, artwork* and *chapter*. These provide besides the *title* and the *source*, a textual *description* and an object with *attributes*. The property of the attribute is provided as the key. The value is an array of objects that itself has a URL and a display value, or simply the value if no URL is available. Figure 17 shows the document-based model that is used for the content types *article* and *youtube*. They provide a *media* object that is an image or a video. Additionally arbitrary content in HTML (*html*) can be included, which will be shown below the media item. For an article this is typically the content of the article itself.

Figure 16: Data model with information about an "entity" item.

Figure 17: Data model with information about an "article" item.

3.4 Content proxy

Within LinkedTV the input for the companion applications comes from the LinkedTV pipeline. The information about the program, such as the title and video stream come directly from the LinkedTV platform. Information about the chapters and dimensions is first curated in the Editor Tool as described in LinkedTV deliverable D1.7. After curation all information that is required by the application about the chapters and dimensions is available in the LinkedTV

platform. For the individual items within the dimensions, however, the application needs additional *display* data. For example, for articles we need the body text (*html*) and the main media item and for artworks the attributes. To get this additional content we developed several services. The basic idea of the service is to resolve the URL and extract the required content from it.

3.4.1 Entity proxy

The entity proxy provides display information for entities from DBpedia. In early LinkedTV prototypes the player directly used the entity proxy. Now it is integrated with the Editor Tool, where the content from the proxy is provided as suggestions to the editors that can select, edit and complement these. The final curated content is then used in the application. The input is a DBpedia URI and the output is a JSON object with display information.

For each DBpedia resource the service tries to return at least a label, a textual description and an image. In addition, the entity proxy can be configured with templates for specific types of entities, for example to return a birth date for persons and the population for cities.

The entity proxy removes the burden of dealing with content from a heterogeneous dataspace such as DBpedia. Typically content for DBpedia resources is collected by either resolving URIs and then getting the right display properties out the response or by submitting a SPARQL query that defines the properties. Getting content in this way is, however, not straightforward.

- DBpedia has a complex schema, where properties are represented in various ways. Dealing with this diversity in SPARQL queries is cumbersome.

- DBPedia contains representations of a resource in multiple languages, represented by different domains: en.dbedia.org, de.dbpedia.org, nl.dpedia.org. Some properties are available for resources in one domain but not in others. For example, the German DBPedia may contain an image of a German minister while the English DBPedia does not. The output of the LinkedTV contains URIs from various DBPedia domains. In all cases we want to return an image when it is available in whatever domain.

Given a DBpedia URI the entity proxy will resolve it and store the RDF response in a triple store. It then looks up in this stored information all the resources that are equivalent. This includes resources from other DBpedia domains, e.g. from en.dbpedia.org to de.dbpedia.org as well as resources that are redirected to another resource. All equivalent sources are also resolved and the RDF response is added to the triple store. The combined information set is now available to get the display properties. The proxy collects the default information, label, description, and image. The RDF properties that can be used to fetch this information are configured in the entity proxy. Finally, resource specific properties are collected. This is based on the type of the entity and the template defined for this type. Both are defined as rules written in Prolog. The entity types are provided as a mapping between a type, e.g. *person* and the URI of and RDF class. In the example below three classes are defined as persons and three others are defined as artists. These mappings help deal with the diversity

where some resources are types as a person by one class and another resource by another class.

entity_type(C, person) :- rdf_equal(C, 'http://dbpedia.org/ontology/Person').

entity_type(C, person) :- rdf_equal(C, 'http://xmlns.com/foaf/0.1/Person').

entity_type(C, person) :- rdf_equal(C, 'http://dbpedia.org/class/yago/Person100007846').

entity_type(C, artist) :- rdf_equal(C, 'http://dbpedia.org/ontology/Artist').

entity_type(C, artist) :- rdf_equal(C, 'http://dbpedia.org/class/yago/Artist109812338').

entity_type(C, artist) :- rdf_equal(C, 'http://dbpedia.org/class/yago/Painter110391653').

The display properties are defined for an entity type. Each property is identified by a name, e.g. *birthDate*, and contains a list of RDF properties used in the data.

entity_prop(person, birthDate,   ['http://dbpedia.org/ontology/birthDate'], false).

entity_prop(person, deathDate,   ['http://dbpedia.org/ontology/deathDate'], false).

entity_prop(person, birthPlace, ['http://dbpedia.org/ontology/birthPlace'], true).

entity_prop(person, deathPlace, ['http://dbpedia.org/ontology/deathPlace'], true).

entity_prop(person, nationality, ['http://dbpedia.org/ontology/nationality'], true).

entity_prop(person, profession, ['http://dbpedia.org/ontology/occupation'], true).

entity_prop(person, predecessor, ['http://dbpedia.org/ontology/predecessor'], true).

entity_prop(person, successor, ['http://dbpedia.org/ontology/successor'], true).

entity_prop(artist, style, ['http://dbpedia.org/ontology/movement'], true).

In addition to the template another advantage of the entity proxy is the reasoning over types and equivalent resources. Reasoning over the type hierarchy means that for a resource of type *artist* it will return the properties defined for this type, but also the properties of type person, as an artist is defined in DBpedia a subclass of a person. Reasoning over equivalences means that properties defined for a resource from the German domain de.dbpedia.org are returned when the input is a resource from the English domain en.dbpedia.org as these resources are defined as equivalent in DBpedia.

The entity proxy accepts a language parameter. If this is provided it will prefer values for the display properties in that language. If these are not found values in English will be returned.

3.4.2 Document proxy

Webpages are an important source for LinkedTV enrichments. The applications for the RBB news as well as the Dutch *Tussen Kunst & Kitsch* program provide a dimension with background information in the form of Webpages. Within LinkedTV these pages are taken from the crawled white lists that were defined for these programs and exposed through the IRAPI web service. In user trials, an editor selected the items for this dimension from the IRAPI service using the Editor tool. The output of the editor tool is the URL of the item and a title. To collect the additional information needed by the application we developed the document proxy. The service takes as input a URL of a Webpage and returns a JSON object with display information. The display information the proxy tries to return is the title, the body text of the Web page, the main media item contained in the webpage, an author name, and an icon representing the source. The document proxy is essential for creating the uniform presentation of content. By extracting specific parts of the Web page, such as the body text and the main media item, the application can repurpose them in a way that best suits the application. Content on the web page that is irrelevant for the application context, such as the page navigation, branding and advertisement are left out.

The document proxy is built around the python library newspaper[4].

3.4.3 API specific wrappers

The document proxy deals with the generic case where enrichments are Webpages. For some enrichment dimensions the content is taken from sources that provide an API. For example, the Related artworks dimension in the LinkedCulture application shows content from Europeana. In the presentation we want to show the attributes of the artwork, such as the creator, style or location. For this case we developed an API specific wrapper that fetches the required content from the record API[5]. It then maps the appropriate properties from the response to the attributes section of the application data model. For other services a similar wrapper can be created.

---

[4]   https://github.com/codelucas/newspaper

[5]   http://europeana.eu/api/v2/record

4   WP3 applications and services

At the end of the LinkedTV project WP3 has delivered: A prototype companion application for TV newscasts; The Multiscreen Toolkit to support the development of multiscreen applications; several prototype application and demonstrators developed on top of the Toolkit; a collection of tools and services to generate display content and a mobile application for an enriched video player.

4.1   LinkedTV News prototype

The LinkedTV News prototype is a second screen application for tablets that acts as a companion to viewers when watching news broadcasts. Its main goal is to enrich news newscasts by integrating them with other media. It is designed to accommodate two viewing modes in terms of interaction: a lean back mode and a lean forward mode.

A video explaining the concept and demonstrating the application is available at:

https://vimeo.com/117467802

Links to the live demonstration are available at http://linkedtv.project.cwi.nl/

4.2   Multiscreen Toolkit

To support the development of applications that are distributed over multiple screens Noterik developed the Multiscreen Toolkit. It supports application developers by abstracting away the low-level details of the synchronization and distribution of content between screens. Developers create a single application that is independent of how many screens are involved and such an application can dynamically react to changes in the amount and types of screens attached to it.   Applications are developed using standard technologies, such as HTML5 and Java.

The toolkit provides input and output services to integrate other LinkedTV components with the applications. The remote API allows external programs to send simple commands to applications, such as play and pause. Through this API a player can be controlled with a remote control or through human gestures. The toolkit provides an output service through the GAIN interface. Through this service other applications can listen to information about the actions that a user performs in the interface. The toolkit is described in the LinkedTV deliverable D5.5.

The Sprinfield Multiscreen Toolkit is available as open source software GNU GENERAL PUBLIC LICENSE Version 3. The software is available at:

http://noterik.github.io/

Noterik is promoting the Multiscreen Toolkit as one of their core products as an addition to their Springfield Framework. The Toolkit is also used in the EU projects Europeana Space[6] and EUscreenXL[7]. Development with the Toolkit is aided by several example applications for which the source is made available on https://github.com/Noterik.

Within LinkedTV the Toolkit is used for the development of several prototype applications. The interfaces are described in LinkedTV deliverable D3.7.

4.2.1 HbbTV single screen application

Noterik, RBB and Condat developed an HbbTV application. This application runs on HbbTV compatible TVs and set-top boxes. The application provides an overlay over a TV broadcast. It provides a straightforward way to access background information about the entities occurring in a TV program. The interaction with the interface is done through the remote control of the TV or set-top box.

A video of the single screen HbbTV application is available at:

https://vimeo.com/106264077

4.2.2 Noterik second screen application

Noterik developed a generic second screen application that synchronizes with the HbbTV application. The application gives access to entities related to the TV program on an additional device. In addition, it allows the user to bookmark and share items.

The source code of the Noterik second screen demo application is available at:

 https://github.com/Noterik/smt_demolinkedtvapp

4.2.3 LinkedNews and LinkedCulture demonstrators

CWI developed the demonstrator LinkedCulture for the AVROTROS program Tussen Kunst & Kitsch" It provides a novel way to experience the program. The user starts the application on the tablet and selects an episode or chapter to watch. Optionally the video is beamed to the TV. When a chapter is finished the user explores rich background and related information about the artwork discussed in the chapter. Finally, the user chooses the next chapter to watch from the suggestions provided by the LinkedTV system.

A video of the LinkedCulture demonstrator is available at:

https://vimeo.com/108891238

CWI developed the demonstrator LinkedNews for the RBB Aktuell News show. The companion application for News uses the same infrastructure as the application for cultural heritage. In this case the enrichment dimensions are tailored to the news domain. In addition

---

[6]   http://www.europeana-space.eu/

[7]   http://www.euscreenxl.eu/

the application provides a second screen mode that is synced with the TV broadcast and shows basic information about the entities that occur in the news.

> The source code of the LinkedNews and LinkedCulture applications is available at:
>
> https://github.com/michielhildebrand/tkk-demo

Links to the live demonstrations of LinkedCulture and LinkedNews are available at: http://linkedtv.project.cwi.nl/

4.3 Content tools and services

During the development of prototypes we encountered the need for tools to generate the display content for these applications. CWI developed several tools and runs several services to fill these needs as described in section 3.4 of this document.

4.3.1 Entity proxy

The entity proxy provides display information for entities from DBpedia. It is used by several LinkedTV prototypes to generate the information cards for the entities. It is also used within the Editor Tool to generate the candidate information about entities that the editor can then curate and extend. The entity proxy is developed in SWI-prolog[8] on top of the ClioPatria RDF Toolkit[9].

> The source code of the entity proxy is available at:
>
> https://github.com/michielhildebrand/linkedtv_entity_proxy

CWI runs an instance of the entity proxy and will keep this running for at least a year after the end of LinkedTV. The configuration of this instance will not be changed, meaning the entity templates remain the same.

> The entity proxy service is available at:
>
> http://linkedtv.project.cwi.nl/explore/entity_proxy

If the entity proxy is to be used by third parties we advise to install their own instance and configure the templates to their needs.

4.3.2 Content proxy

To generate the content for the LinkedTV companion applications CWI developed a set of tools that are bundled together in the so-called content proxy. The source code is made available at https://github.com/michielhildebrand/linkedtv_content_proxy. The code is not intended for reuse as such, it serves as an illustration for the type of operations that are needed to generate display information.

---

[8]   http://www.swi-prolog.org/

[9]   http://cliopatria.swi-prolog.org/

4.4 Mobile application for enriched video

For the final LinkedTV user trials CWI re-developed the LinkedCulture and LinkedNews demonstrators into standalone mobile applications. The result is an HTML5 application that can be configured for a specific TV program by feeding it data according to the data model described in Section 3.3 of this document. The application is built on top of the Ionic HTML5 hybrid mobile App framework[10]. The application can be deployed as a standard Web application that is accessed in a Web browser or build into an Apple iOS or Android application. Links to the demonstrators used in the user trials are available from http://linkedtv.project.cwi.nl/.

CWI is setting up a valorization project in combination with the CWI spinoff Spinque[11] to further develop this application and apply it to different domains. Spinque has started development of a demonstrator with the application within the EU project Comsode[12].

---

[10]   http://ionicframework.com/

[11]   http://www.spinque.com/

[12]   http://www.comsode.eu/

5   Bibliography

[1]   David Milne and Ian H. Witten. 2008. Learning to link with Wikipedia. In *Proceedings of the 17th ACM conference on Information and knowledge management* (CIKM '08). ACM, New York, NY, USA, 509-518.

[2]   Vadim von Brzeski, Utku Irmak, and Reiner Kraft. 2007. Leveraging context in user-centric entity detection systems. In *Proceedings of the sixteenth ACM conference on Conference on information and knowledge management* (CIKM '07). ACM, New York, USA.

[3]   Perez Romero, L., LinkedTV News designing a second screen companion for web-enriched news broadcasts. Eindhoven University of Technology, 2013.

[4]   Hauptmann, A.G. and Witbrock, M.J. Informedia: News-on-Demand Multimedia Information Acquisition and Retrieval. Intelligent Multimedia Information Retrieval. AAAI Press, 1997, 213-239.

[5]   Mike Dowman, Valentin Tablan, Hamish Cunningham, and Borislav Popov. 2005. Web-assisted annotation, semantic indexing and search of television and radio news. In *Proceedings of the 14th international conference on World Wide Web* (WWW '05). ACM, New York, NY, USA, 225-234.

[6]   M. Dowman, V. Tablan, H. Cunningham and B. Popov. Content Augmentation for Mixed-Mode News Broadcasts. 3rd European Conference on Interactive Television: User Centred ITV Systems, Programmes and Applications. Aalborg University, Denmark, 2005.

[7]   Przybocki, M., Fiscus, J., Garofolo, J. and Pallett, D., 1999: 1998 HUB-4 information extraction evaluation. In Proceedings of the DARPA Broadcast News Workshop (Herndon, VA), 13-18.

[8]   Maybury, M., 2003: Broadcast News Understanding and Navigation. In proceedings of Innovative Applications of Artificial Intelligence (Acapulco, Mexico), 117-122.

[9]   Daan Odijk, Edgar Meij, and Maarten de Rijke. 2013. Feeding the second screen: semantic linking based on subtitles. In *Proceedings of the 10th Conference on Open Research Areas in Information Retrieval* (OAIR '13), João Ferreira, João Magalhães, and Pável Calado (Eds.). LE CENTRE DE HAUTES ETUDES INTERNATIONALES D'INFORMATIQUE DOCUMENTAIRE, Paris, France, France, 9-16.

[10] Monika Henzinger, Bay-Wei Chang, Brian Milch, and Sergey Brin. 2003. Query-free news search. In *Proceedings of the 12th international conference on World Wide Web* (WWW '03). ACM, New York, NY, USA, 1-10.

[11] Ardissono, L., Kobsa, A., and Maybury, M., (eds.) 2004. Personalized Digital Television: Targeting Programs to Individual Users. Springer Verlag, London.

[12] Lindstedt, I., Löwgren, J., Reimer, B., and Topgaard, R. 2009. Nonlinear news production and consumption: A collaborative approach. ACM Comput. Entertain. 7, 3.

[13] Olsen, D. R., Bunn, D., Boulter, T., and Walz, R. 2012. Interactive television news. ACM Trans. Multimedia Comput. Commun. Appl. 8, 2.

[14] Zimmerman, J., Dimitrova, N., Agnihotri, L., Janeyski, A., and Nikolovska, L. Interface Design for MyInfo: a Personal News Demonstrator Combining Web and TV Content. INTERACT, IOS Press, (2003).

[15]   LinkedTV News: A dual mode second screen companion for web-enriched news broadcasts 2014 - Proceedings of ACM International Conference on Interactive Experiences for Television and Online Video 2014 (ACM TVX 1), Newcastle, UK (demo track)

[16]   J. Sivic and A. Zisserman. Efficient visual search for objects in videos. Proc. IEEE, 96(4):548–566, 2008.

[17]   Khaoula Elagouni, Christophe Garcia, and Pascale Sébillot. 2011. A comprehensive neural-based approach for text recognition in videos using natural language processing. In *Proceedings of the 1st ACM International Conference on Multimedia Retrieval* (ICMR '11). ACM, New York, NY, USA, Article 23 , 8 pages.

[18]   Dowman, M.; Tablan, V.; Cunningham, H.; Ursu, C. & Popov, B. (2005), Semantically Enhanced Television News through Web and Video Integration, *in* 'Second European Semantic Web Conference (ESWC 2005)'.

[19]   Redondo García J.L., Hildebrand M., Perez Romero L., Troncy R., Augmenting TV Newscasts via Entity Expansion. In ESWC 2014, Posters and Demos Track May 2014, Anissaras, Hersonissou, Crete.

[20]   Lotte Belice Baltussen, Roeland Ordelman, Jaap Blom. VideoHypE: An Editor Tool for Supervised Automatic Video Hyperlinking. In *11th European Interactive TV Conference (EuroITV 2013)*

[21]   Journalism website. Press release: getting the facts straight. Work by Owen Spencer-Thomas, D.Litt. URL retrieved 25 February 2015.