# Television linked to the Web

Proposal full title:
Television linked to the Web

Proposal acronym:
LinkedTV

Type of funding scheme:
Integrated Project (IP)

Work programme topics addressed:
Challenge 1:     Objective ICT - 2011.1.5: Networked Media and Search Systems
Target outcome a) Digital Media Delivery Platforms

Name of the coordinating person:        Dr Joachim Köhler, Fraunhofer IAIS
e-mail:                                  joachim.koehler@iais.fraunhofer.de
fax:                                     +49 2241 14 2342

List of participants:

| Particip ant no. * | Participant organisation name | Part. short name | Country | Type |
|---|---|---|---|---|
| 1 | Fraunhofer IAIS | FRAUN-HOFER | Germany | Public Research Center |
| 2 | Centre for Research and Technology Hellas | CERTH | Greece | Public Research Center |
| 3 | University of Economics, Prague | UEP | Czech Rep | University |
| 4 | Centrum Wiskunde & Informatica | CWI | Holland | Public Research Center |
| 5 | Eurecom | EURE-COM | France | Public Research Center |
| 6 | University of St Gallen | UNISG | Switzerland | University |
| 7 | Semantic Technology Institute International | STI | Austria | Public Research Center |
| 8 | Condat | CONDAT | Germany | SME |
| 9 | Noterik | NOTERIK | Holland | SME |
| 10 | Sound and Vision | S+V | Holland | Public Non-profit |
| 11 | University of Mons | UMONS | Belgium | University |
| 12 | Rundfunk Berlin Brandenburg | RBB | Germany | Public Broadcaster |

# Abstract

*LinkedTV is an integrated and practical approach*
*towards experiencing Networked Media in the Future Internet!*

Networked Media will be a central element of the Next Generation Internet. Online multimedia content is rapidly increasing in scale and ubiquity, yet today it remains largely still unstructured and unconnected from related media of other forms or from other sources.

This cannot be clearer than in the current state of the Digital TV market. The full promise and potential of Web and TV convergence is not reflected in offerings which place the viewer into an Internet closed garden, or expect PC-like browsing on a full screen Web, or offer interesting new functionalities which however lack any relation to the current TV programme.

Our vision of future **Television Linked To The Web (LinkedTV)** is of a ubiquitously online cloud of Networked Audio-Visual Content decoupled from place, device or source. Accessing audio-visual programming will be "TV" regardless whether it is seen on a TV set, smartphone, tablet or personal computing device, regardless of whether it is coming from a traditional or new media broadcaster, a Web video portal or a user-sourced media platform.

Television existing in the same ecosystem as the Web means that television content and Web content should and can be seamlessly connected, and browsing TV and Web content should be so smooth and interrelated that in the end even "surfing the Web" or "watching TV" will become as meaningless a distinction as whether the film is coming live from your local broadcaster, as VOD from another broadcaster, or from an online video streaming service like Netflix.

As a result, not only commercial opportunities but also opportunities for education, exploration and strengthening European society and cultural heritage arise. Imagine browsing from your local news to Open Government Data about the referenced location to see voting patterns or crime statistics, or learning more about animals and plants shown in the currently viewed nature documentary without leaving that show, or jumping from the fictional film to the painting the character just mentioned to virtually visiting the museum when it can be seen, or seamlessly accessing additional information that has been automatically aggregated from multiple sources in order to get better informed on an important event that was just mentioned in the news.

Technologically, this vision requires systems to be able to provide networked audio-video information usable in the same way as text based information is used today in the original Web: interlinked with each other at different granularities, with any other kind of information, searchable, and accessible everywhere and at every time. Ultimately, this means creating hypermedia at the level of the Web.

The Web's original success was the underlying hypertext paradigm built into HTML. Hypermedia has been pursued for quite a while as an extension of the hypertext approach towards video information. But it needs complex video analysis algorithms and is still an issue of research. **Television Linked To The Web (LinkedTV)** provides a novel *practical* approach to Future Networked Media. It is based on four phases: annotation, interlinking, search, and usage (including personalization, filtering, etc.). The key element is a new practical approach to annotation by using a broad spectrum of information available in the media production process, on the Web, and from users for video content annotation: automatic annotation from picture and video sequence analysis, related audio information, scripts, production plans, Wikipedia, geo-spatial information, manual annotations by content providers and/or users. We aim at a fine-grained concept-, instance- and event-based

annotation not just of video as a whole but also of its relevant parts. Multimedia information annotated in this way can be *interlinked* to each other at different granularities and to other kinds of information. It can be *searched* on a detailed level, and be adapted to users' personal needs, concrete contexts, and preferences.

Hence we will provide the technology to enable a new generation of television in line with the shift of content to the rapidly ubiquitous Internet and the blurring lines between Web and Television: Television Linked To The Web, in which we seamlessly weave multimedia content across the Internet by integrating networked media analysis, personalisation and presentation technologies within an integrated and coherent framework. LinkedTV will develop a comprehensive methodology, as the basis for the LinkedTV platform and tools. This platform will enable the information management and usage in the annotation, interlinking, and personalization process. It supports the media production process in a novel and integrated way, and provides support to users by providing linked information, filtering for personal preferences, search capabilities, etc. The LinkedTV platform will support the whole Networked Media process from end to end in an integrated and coherent way. This will be provided by building on existing and emerging standards such as HTML5, MPEG7, MPEG21 and SMIL as well as the results of past and current research projects on context analysis and representation, multimedia analysis, annotation, adaptation and delivery. LinkedTV will extend available tools, fill current gaps and create needed connections in order to support video links seamlessly for the end-to-end process.

The result will make Networked Media more useful and valuable, and it will open completely new areas of application for Networked Audi-Visual Content, defining a new experience of media access and interaction that we believe will be part of the European citizen's experience of Television in the Future Internet. Since Television Linked To The Web will become a core technology for the network of the future, we include business modelling and industrial exploitation partners who will drive market entry of this disruptive television experience, forming the basis for a new generation of advanced Networked Media applications. LinkedTV will provide the technologies needed in the broadcast and telecommunication industries to enable innovative Networked Media services which can create new market opportunities for SMEs and protect and improve the competitiveness of large traditional market players in an increasingly fractured future media landscape. LinkedTV demonstrators will be used to present Television Linked To The Web to the public (end user) as well as to industry players in the broadcast, telecommunication and IP hardware industries (technology adopters). Exploitation will be assured by timely market surveys, value chain analysis and preparation of appropriate business models.

# Executive Summary

There are still important gaps and breaks in the current "Networked Media Web". TV-over-Web does not yet make full use of the interaction possibilities inherent in the Web model, choosing rather to shift the passive watching experience to the screen of a computer (or smartphone). Web-over-TV is either about putting a Web browser onto the TV screen or overlaying parts of the Web over the TV signal; neither explore how one could browse between TV and Web content in a richer and more seamless manner. The vision of the realisation of a truly Networked Media Web is much more challenging: it means to provide audio-video information on the Web usable in a way similar as text based information is used today: interlinked with each other, with other kinds of information, searchable, and accessible everywhere and at every time.

Hypervideo has been pursued for quite a while as an extension of the hypertext approach towards video information. But it needs complex video analysis algorithms and is still an issue of research. **Television Linked To The Web (LinkedTV)** aims to provide a novel *practical* approach to Networked Media based on four phases: (1) annotation, (2) interlinking, (3) retrieval, and (4) presentation (including personalization, filtering, etc.).

LinkedTV will allow us to seamlessly connect multimedia content on the Web by integrating networked media analysis, personalisation and presentation technologies within an integrated and coherent framework. LinkedTV will develop a comprehensive, coherent methodology for the end-to-end process, as the basis for the LinkedTV platform and tools. This platform will enable the information management and usage in the annotation, interlinking, and personalization process. It supports the media production process in a novel and integrated way, and provides support to users by providing linked information, filtering for personal preferences, search capabilities, etc.. This will be supported by building on the results of FP6 and FP7 projects on context analysis and representation, multimedia analysis, annotation, adaptation and delivery.

The result will make Networked Media more useful and valuable, and it will open completely new areas of application for Multimedia information on the Web.

# Section 1. Project Description

## 1.1 Concept and objectives

The World Wide Web already today has completely changed the way we are providing and using information. It has grown from a static, text and image based information medium to a much more dynamic, interactive, multimedia, and fixed-and-mobile network. Video content on portals like YouTube is growing in amount and consumption at an incredibly fast rate, spreading out from the PC onto the television set and mobile devices. It is clear that the future Internet will carry much higher amounts of video material than even today, and will become the dominant source of audiovisual information. Also classical television is being absorbed into IP networks as both live streams and catch-up, alongside other providers of video, and the content is mixed at the end device blurring the line between "Web" and "TV" content. However rich interlinking between this content is lacking, making TV-Web convergence offers today still very limited compared to the richness of browsing the World Wide Web.

This vision of the Networked Media Web is challenging: it means to provide audio-video information on the Web usable in a way similar as text based information is provided today in the Web browser: interlinked with each other, with other kinds of information, searchable, and accessible everywhere and every time. Just as hypertext fulfilled its potential in the Web through a set of related standards (HTTP, HTML, URLs), hypervideo – Television linked to the Web – can not fulfil its potential without standards and technologies. Many of the building blocks exist or are being formed today (HTML5, Linked Data, a common Media Ontology, Media Fragment identification) while some aspects are missing or incomplete (shareable user profiles, concept-level video analysis, dynamically linking media segments to related content, abstract presentation templating). LinkedTV will provide a tool chain seamlessly connecting these standards and technologies into an end-to-end platform for interlinking TV and Web content and enabling a new experience of Networked Media. Three pilots will be built to validate, elaborate and push current standards and technologies, forming the basis for the industry maturity and uptake of LinkedTV.

*From the Web and hypertext to LinkedTV and hypervideo*

Two features of the mainly text-based Web made it as accessible and usable as it is today: hypertext and search engines. Hypertext allows content creators to interlink related pieces of information. Search engines allow people to search what they need in the huge amount of information available on the Web. Both features are not available today in the same sense for rich media content such as video. The content of video is not transparent to search engines, and interlinking has to be done manually for complete videos. In order to be interlinked with their content and to be searchable they need annotations. Today, videos can be annotated by their creators or in social networks - but as a whole and on a keyword and simple metadata level. These annotations can be used by classical text-based search and for interlinking.

Television Linked To The Web is a central element of the vision for the future Networked Media Web. It will bring the last remaining established paradigm for consumption of video outside of the Internet fully within the rich interlinked model of Web content. This will in turn change the way we use audio-visual information in much the same way as the Web has changed the way we use textual information. The human visual sense is our most powerful "information channel" to our outside world. Video information will be used in much more cases than today if it can be accessed in the same way as text on the Web: ubiquitously available, easy to find, and dedicated to the special purpose at hand. Why read boring manuals if a short video sequence allows people to find out what they need? Why read long articles about something if it can be explained in a video? Or if video information can augment and illustrate text based information? Furthermore, content can be interlinked to another video going more into the details, explaining a related issue, etc. complemented by

links to other media such as texts, graphics, and audio presented where relevant in an appropriate and informative way (e.g. pictures of an event arranged chronologically).

Hypervideo is a research topic which aims at an extension of the hypertext approach towards video information. But it needs complex video analysis algorithms and is still an issue of research. LinkedTV will pursue a *practical* approach towards linking television into the Web which is now manageable. It will aim at a fine-grained annotation of video content using a broad spectrum of methods and information. Not just the whole video will be annotated but each relevant piece of its content on a level of detail allowing people and systems to use it in a dedicated way and to search for it. This fine grained annotation enables the rich interlinking of video content with all kinds of Web based information - including all the other videos. Full video annotation including automatic picture and video sequence analysis will continue to be a research issue for quite a while. LinkedTV will use as much of automatic and semi-automatic picture and video analysis and annotation as feasible (including results from other European research projects). *Additionally*, all other kinds of information available in the process of content creation and from other sources will be used for annotation in a systematic and integrated manner: production scripts, subtitles, audio information, Web based information like Wikipedia, geo-information, etc. This integrated approach needs an elaborated methodology how to treat the many different kinds of information in a systematic way.

*Goal and vision of LinkedTV*

Consequently, LinkedTV will concentrate on three tightly interconnected main goals towards a practical approach to a Networked Media Web:

- LinkedTV will develop a comprehensive methodology for information management to be used in the course of fine grained video annotation. This provides the basis for interlinking, search, personalization, and contextualization.

- We will create the LinkedTV platform in order to demonstrate the feasibility of our approach including all necessary components from annotation, interlinking, and search to personalization and effective, efficient and intuitive usage by users.

- Three representative usage scenarios of quite different types spanning a broad range of potential use cases of our approach will be developed as proof of concept.

We expect that LinkedTV will become the basis for a wide range of advanced hypervideo applications for different contexts beyond those which we will prototypically consider (PC and TV-based access), including mobile, 3D Internet and virtual worlds. The hypertext-centered Web will be extended by a hypervideo-centered Television Linked To The Web, where the access paradigm is not the HTML-based browser but an audio-visual channel from which the viewer can access concepts and browse content.

*LinkedTV innovation, output and benefits*

To summarize, here are the problem areas addressed by LinkedTV, the innovations of this project, the output foreseen in the work programme and the benefit to users:

**Table 1: Problem Areas**

| Problem Area | Innovation | Output | User benefits |
|---|---|---|---|
| Current audiovisual analysis approaches do not provide fine | Novel semi- or fully-automatic analysis algorithms to assist in | Tool for concept-level annotation of fragments of | Broadcasters can more efficiently and effectively annotate |

| | | | |
|---|---|---|---|
| grained enough identification of individual concepts in the material | annotating content; interaction with the user to allow efficient search and identification of similar regions or events across media. ; better automated identification of concepts through complementary material analysis | audiovisual material (WP1), incorporating conceptual annotation support through complementary Web data mining (WP2) | their content. |
| Current multimedia annotation schemes are not sufficient for unambiguous retrieval and packaging of associated Web content | New annotation schemes for Networked Media extending and re-using existing standards and best practices | Specifications for the reference of media fragments and for the annotation of media in terms of conceptual representation (WP2) which will be pushed to standardization | Content owners, broadcasters and providers of added value services can share and re-use annotations |
| Current multimedia services over Web or TV use fixed, manually created interfaces for user interaction, and rely on the effortful preparation of content packages in advance | New interface and content presentation specifications will abstract from their final form, allowing for user and context personalization; multimedia presentations around selected concepts will be generated on the fly | A presentation engine for the on-the-fly construction of user interfaces and multimedia presentations (WP3), an adaptation engine for the dynamic personalization of interfaces/presentatio ns to a given user/context profile (WP4) | Providers of added value services can offer a wider range of new innovative services at lower cost (re-using Web content) and with greater personalization (increasing potential commercialization). |
| Current access paradigms to audiovisual material, including online video or IP-based television, do not present a tight integration with associated content in global networks | New services will be enabled where consumers of audiovisual material may have instant access to information about concepts represented in the viewed material | A LinkedTV platform will integrate all components produced in the project (WP5) and provide the basis for the implementation of individual scenarios (WP6) | Broadcasters can win new viewers and larger market share through exclusive use of content in services. Viewers have access to new added value services in association with TV/video viewing |
| Current efforts at bringing interactive media into mainstream markets have not (yet) been successful (of course, none of these provides what LinkedTV will provide) | In the project, a specific work package shall be dedicated to addressing the issues of exploitation, including the development of business plans, to ensure that project results can be | In particular, the development of the interfaces will be carried out closely with the end user (WP3). The scenarios will be trialed with users (WP6). A dedicated effort will be made on developing business | Potential purchasers and users of the technology are supported in their investment decisions. Potential enterpreneurs and innovators are supported in their business plans and funding applications. |

| | converted into enterpreneurship and new market opportunities | plans and promoting exploitation of project results (WP8). | |
|---|---|---|---|

### 1.1.1 Project description

LinkedTV is about researching and developing novel techniques for connecting heterogeneous Networked Media, therefore making audio-visual content more usable in a much broader field of applications. This will create a wide spectrum of new market opportunities. LinkedTV will offer a practical approach to a completely new usage of video based information on the Web (see Figure 1). We will enable people to search and access Networked Media and related Web content intuitively and instantaneously from within audio-visual material.
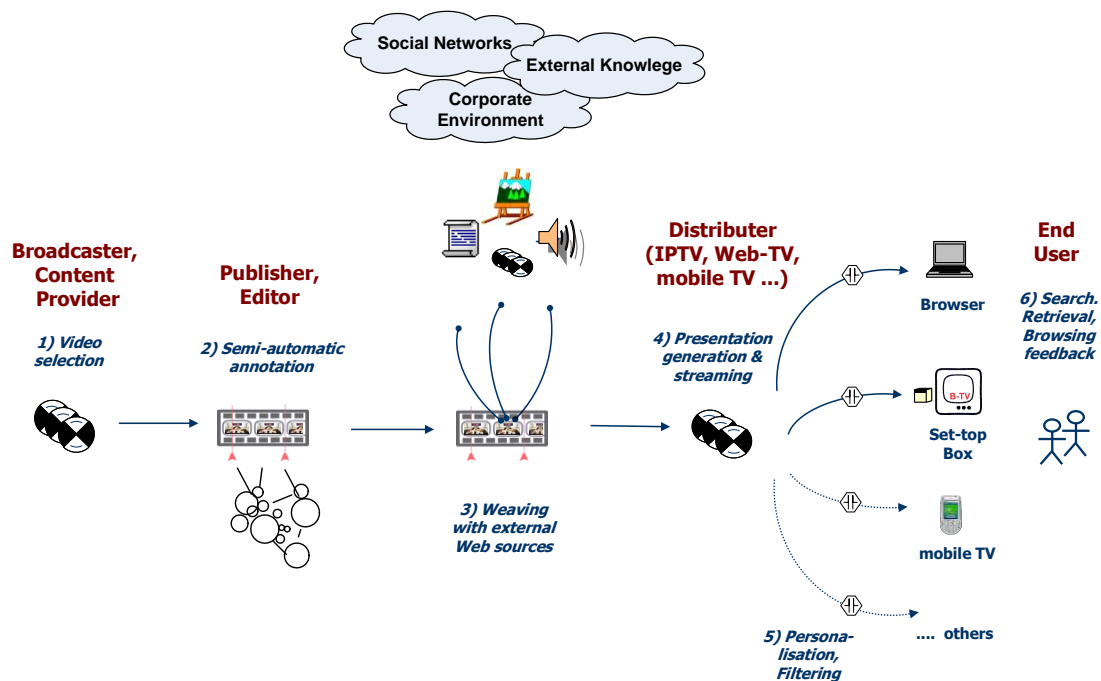


**Figure 1: The core steps of creating television linked to the Web**

Basic Concept :

Audio-visual information can be attractive and effective in many cases: from typical industrial and informational use cases like process or system descriptions to education, training, and all kinds of entertainment. Frequently, a combination of different media forms will be the most effective one: audio-visual information connected to each other, to text, graphics, etc. The problem is that the audio-visual material is not treated and not available on the current Web in the same way as other types of information like texts and graphics are: not at an appropriate granularity, not related to other content, and not sufficiently searchable and accessible.

But even if audio-visual information can be treated on the Web in a similar way as text based information this is not sufficient: we need personalized information and context specific search and retrieval. We need new interaction capabilities of users to make use of the

networked media information in an effective and efficient way – adapted to the many different kinds of usage of audio-visual information (see Figure 2).

Viewers may want to find out the following:



**Figure 2: Sample LinkedTV-enhanced audio-video information**

As a consequence, LinkedTV will address the following main and interrelated four issues:

- Fine-grained annotation of video information using an integrated combination of automatic, semi-automatic, and manual annotation procedures. For this purpose, a broad spectrum of different kinds of information will be used: from picture analysis in video frames over audio information associated to the video, production scripts and documentation, manual annotations by content providers and/or users, etc.

- The LinkedTV methodology and the platform to be developed will support this media production and annotation process systematically and comprehensively. This broad availability of information related to the video from the very first moment of idea creation till the end by consumption will substantially help to run collaborative media production processes more efficiently and to improve their quality and usability.

- Audio-visual information annotated in this way can be interlinked to any other kind of related information on the Web – in whatever form (see Figure 3). Videos can be linked not just as a whole but on a much more fine-grained level to other videos, to graphics, text, geo-spatial information, etc. Annotated in this way they can be searched similarly as any other kind of information.

- The user needs sophisticated support in order to manage this complex networked media world. Depending on the domain of usage (professional and business, education and training, entertainment and social interaction) quite different kinds of support are necessary. Personalization, filtering, and context sensitive search are essential elements. User interactions in such rich and complex media worlds need sophisticated, intuitive, and easy-to-use interaction capabilities.
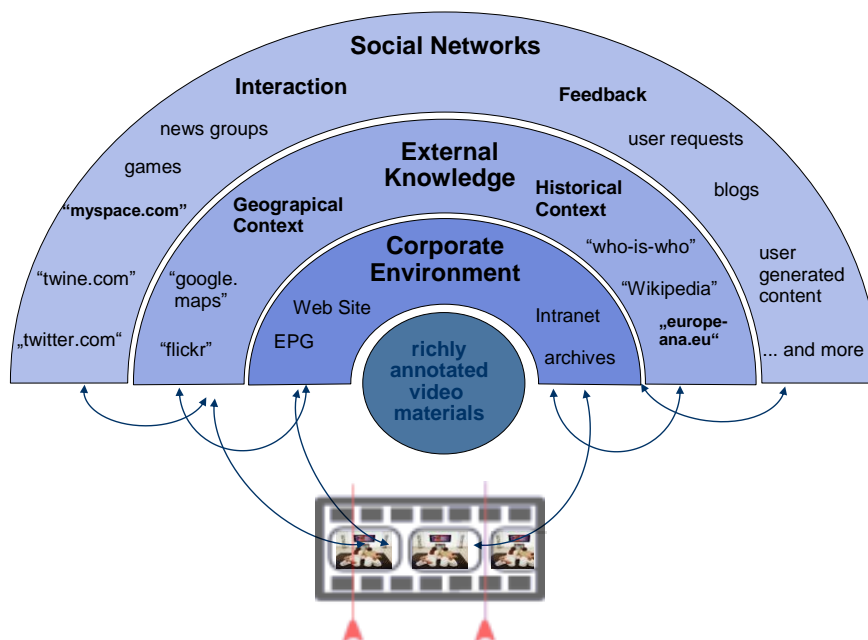
**Figure 3: Weaving Television into the Web**

### 1.1.2 Scenarios

We will demonstrate our methodology of weaving of television into the Web and the capabilities of our LinkedTV platform through three scenarios, each of which representing different aspects of the value and potential of the future Networked Media Web (see also chapter 1.2.1 for a more detailed description):

An **environmental** scenario will showcase how Connected Media Experiences can empower opinion making through instant information access, by retrieving factual information, different opinions and objective evidence. Depending on the individual interest, the user can decide how much he wants to know about an issue. This ensures that people can get information on exactly the right level that matches their abilities, needs and knowledge.

A **documentary** scenario will showcase the value in enabling reuse of digital cultural heritage archives in a new context, putting shared cultural knowledge at the fingertips of every (connected) citizen. We will draw on the rich European media and content archives of Europeana and its video arm, VideoActive.

While the "Documentary" showcase focuses on the user possibilities, the **media art** scenario introduces the artist as a mediator between a public and available multimedia content. Availability of tools for hyperlinking and annotating various forms of content, and the content itself, will constitute tremendous assets in increasing productivity, in providing a platform for experimenting novel approaches in storytelling, or more generally in stimulating creativity. In a context of reuse and repurposing, an artist can browse content in order to rapidly find the data he needs in order to create new content, or even an artistic performance. Indeed, several experiments of live content browsing seen as an art or entertainment form can be found, including the art of skilled live music remixing. Recently, richer content is experimented live by artists, and content browsers are part of the performance, facilitating retrieval, or supporting the storytelling, in settings where the screen is part of a hybrid stage.

Public feedback is an important facet here. In a world of networked creation and distribution, the artist can also benefit from live feedback of explicit (specific gestures) or implicit

(individual emotions, social interactions) behavior of the public. The content can be distributed on hybrid internet-broadast services to classical setups in a home, or in more or less small cinema rooms where a more important crowd can watch and the social interaction can be more complex. Depending on this feedback, the artist can decide to modify his performance or the provided content, the public hence impacting on the interactive story that the artist mediator conveys.

All three scenarios provide quite different requirements to the LinkedTV methodology, the production platform, and the user interface. The quality and granularity of video annotation, the kind of procedures and additional information available and usable, the search capabilities, and the user interaction needs are quite different.

The environmental scenario will frequently need precise and fine grained annotations, associated with content and time-aware information. The LinkedTV platform can make use of all information available in the course of media production including the general ecological information and opinions of the most relevant opinion leaders, bodies and representatives. The LinkedTV user interface will tightly integrate special search functionalities to empower the user to easily identify information sources and classify the content.

The documentary scenario has special requirements to the kind of video content annotation. A rich environment is needed in order to relate the basic annotations to the many different aspects this annotation may have in different cultural contexts. The content of an artistic documentary may be described in many different ways as well as its style. Semantic search is needed to match user queries with these many different views and contexts. The LinkedTV user interface should support the user to manage the rich information associated to cultural content.

The media art scenario may not need so much annotation precision, relying potentially on more serendipitious links between content to create an entertaining experience of content discovery and exploration. It needs sophisticated interaction capabilities and attractive and fancy user interfaces allowing people to make use of the annotations and links in a rich and sometimes unforeseen way.

These scenarios arise out of the recognition that audio-visual information is not an isolated medium in a Networked Media world, but a key node in this complex information landscape. Today, professional content providers like broadcasting organisations have an exclusive focus on producing radio and television, and on publishing program guides. On our way to a Networked Media world they will become cross-media production companies. Audio-visual content will completely be integrated into the Networked Media Web (including mobile, home and corporate platforms). It has to be interlinked with other audio-video information, text based information, etc. much in the same way as hypertext is used today on the Web. Television and other audio-visual archives provide rich content to be seamlessly integrated, accessible and searchable.

Modern audio-visual archives, for instance in television broadcasting organisations, cater for a broad range of users who display a broad palette of search behaviours. Broadcast professionals often have very focused information needs expressed using a mixture of textual and metadata queries. Media researchers such as documentary makers or television historians, on the other hand, often mix search and undirected browsing, jumping to related items so as to explore an entity, theme or event. Historians, for example, rarely search by subject headings [1, 2]. What these users share is a need to see and access items in their broader context, as well as a strong orientation towards entities (e.g., people, locations, artefacts, etc.), themes (e.g., "festivities," "consumer culture," "catastrophes," "living in cities," etc.), and events ("the Prague spring of 1968," "the opening of the channel tunnel in 1994," etc.) [3, 4]. Hence, meaningful access to modern audio-visual archives has to allow search and discovery by entities, themes, and events, and by the relationships between them and a specific context.

At this point in time, audiovisual archives (such as those represented by partner Sound and Vision) are rapidly complementing traditionally archived materials with additional materials (user generated content, crawls of broadcast related websites, program guides, etc). For instance, with the help of the EU the Europeana portal was built. (www.europeana.eu), giving users direct access to over 5 million heterogeneous digital objects [5]. On the Web, sources like Wikipedia, but also user generated content on YouTube, Flickr and other "Web 2.0" sites can provide meaningful media content related to a given programme. This is being extended by a metadata layer on the Web based on so-called "Semantic Web" technologies which can enable concept-based search and browsing for information. All of this represents relevant content which may be meaningfully associated with an audiovisual broadcast (Figure 4).
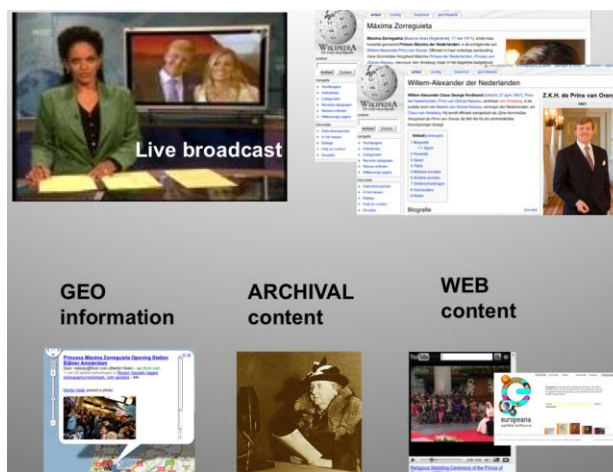


**Figure 4: Sample Components of Television linked to the Web**

### 1.1.2.1 Scenario 1: Using audio-video content in environmental information

Paul has recently moved to Brandenburg and is now adapting to his new neighbourhood. A few days ago, he was to asked sign a petition against a new wind turbine park. He refused because he knew nothing about the project. Today he is sitting in his living room and is watching the local TV-station "Rundfunk Berlin-Brandenburg" (rbb). Currently, the science TV-Show "OZON unterwegs" is on. It is reporting on wind power generation in the lowlands of Brandenburg. This arouses his interest. The companies behind the shown wind turbines and the opponents of the project are of particular interest to him. He would like to see a short history and profile of the company (text-based) as well as recent news material that mentioned the company (text/audio/video) which should be presented in a chronological manner. Paul also wants to find out about the political situation and context regarding environmental and economical issues. Ideally, Paul wants to obtain all this information as objectively and efficiently as possible in order to be able to make an informed, facts-based decision whether he should sign the petition. In addition to the above, Paul also wants to evaluate the popularity of ecological measures by analyzing user-generated content surrounding the service in question. A multimedia service that offers him a great variety of high quality multimedia content from different sources presented in an intuitive way, based on his specific needs and requirements, would be a highly appealing service for Andy. LinkedTV is to provide just that!

John is an engineer working for a company that produces wind power stations. Sitting at his desk, he is watching a program about a wind power plant. John is interested in the mechanical aspects of the turbines and would like to access images/models of the actual turbine, presentations and videos of how it operates, and text with further supportive information (e.g. details of the turbine's features, components, manufacturer etc). He has the technical

background knowledge to recognize advantages and disadvantages of the technology deployed and uses the service that is to be developed in LinkedTV as a thought-provoking impulse for his work. By using the service, he also stays up-to-date with technological developments, and stays informed about the political and economical situation about the industrial sector, which is very important for him.

Laura has just finished school and is now considering what she would like to study. At her PC she is watching a program on wind turbines. She wants to understand (on a basic level) how they function and what she would have to study if she wants to become a wind turbine engineer. She furthermore wants to find out where she could study this and what the future prospects are.

Being an ecologically aware person, Laura is also interested in saving money by acting economically with regard to power consumption, heating costs and fuel consumption of her car. She wants to compare her power consumption patterns to that of households in her neighbourhood, with other people who are in a comparable situation, and do so both on a regional as well as national level. The tools and services to be developed in LinkedTV, together with already existing third party services (e.g. platforms that compare prices such as www.verivox.de) help her find the best offers that meet her private needs and agree with her ecological conscience. Through LinkedTV, Laura is in a position to make more informed decisions.

This environment scenario is characterized by the need of precise and fine grained annotations. For the content producers the LinkedTV platform is very attractive to represent their opinion for interested. The LinkedTV platform can support the content creators with all information available in the course of media production including the general business and technical information related to the case at hand. For users like Paul, John, and Laura the search functions will need high precision.



| Concept | Diana Brätz | A Wind Turbine |
|---|---|---|
| Class | Person, Local Resident | Wind Turbine |
| Related content | **Local residents**<br>Information about the campaign of the local residents<br>Campaign web site | **General information:**<br>Description „what is a wind turbine?"<br>Diagrams „how does it work?"<br>Images/videos of turbines<br><br>**For an expert (e.g. an engineer):**<br>Turbine specifications<br>Diagrams<br>Technical manuals |

| Browsing facets | **To the wider media sources**<br>Related media coverage about the campaign<br><br>**To other opinions** | **To general information**<br><br>**To the political context** |
|---|---|---|

**Figure 5: Example for the environment scenario**

### 1.1.2.2 Scenario 2: Documentary

The envisaged documentary use case will use the cultural heritage domain to equip researchers with the following facilities: (i) starting from a description, identify relevant related items and background information; (ii) aggregation of information about entities being mentioned in the descriptions and related sources; (iii) presenting the resources in an appealing way.

Ruth is working on a student project about musical instruments in the eighteenth century. She goes online and soon finds a documentary from the Sound and Vision on Mozart. It shows Vienna and the interior of the Mozart museum. An eighteenth-century piano is on display. The LinkedTV system makes the system aware that more information is related to this object. By clicking on the item, paintings that include similar piano's (available on www.european.eu) are shown, next to a segment of another documentary on the company that made musical instruments in Salzburg at that time. Ruth chooses to look at the painting more closely and examines the information available about it. It turns out that the painting is made by the French painter Luis Carrogis and Wikipedia content about Carrogis is presented subsequently, alongside a representation of the Getty ULAN thesaurus that shows his relationship with other people at that time. It turns out he offered his services to Napoleon. By clicking on Napoleon, LinkedTV offers access to a documentary on the life of Napoleon, again from Sound and Vision. Using the transcripts, Ruth searches for "music" and is directed to the segment that deals with marching music. Here, she notices a typical type of drums. In an overlay, LinkedTV indicates that this type is part of the Louvre collection on musical instruments.

Julien and Thomas are used to watching the Jeugdjournaal, the daily Dutch news programme targeted primarily at youths, at 7 pm. For a few weeks the interactive version of the Jeugdjournaal has been offering LinkedTV functionality, a service the two brothers like a lot more than the 'traditional' version. It provides an interface that shows archived content alongside the news broadcast. The collection of archive content originates from broadcasters and news agencies and comprises material of different durations, origins and temporal provenance. Alongside every segment in their news broadcast this contextual material is offered. Julien can choose different modes for recommendation, namely who, what, where and when. He starts watching the latest developments on the war in Afghanistan in the 'What' mode. At a certain point, UN soldiers are on the screen, wearing blue helmets. LinkedTV subsequently offers Wikipedia content about UN peace keeping troops and a short clip from Sound and Vision showing the UN peace troops on earlier assignments. The next item of the Jeugdjournaal is dealing with a fire in Beijing. Julien switches to 'Where' mode and receives information on the city of Beijing, pictures from Flickr, statistics from The CIA World Fact Book, and a clip made in 1998 covering the city's bid for the 2008 Olympic Games. One of the shots shows the Niujie Mosque. Thomas forwards a reference to this video to his friend Walter who has an interest in mosques.

**Figure 6: Content used in the documentary scenario**



| Concept | UN Peace Troops | French flag emblem |
|---|---|---|
| Class | Soldiers | Flags |
| Related content | **Online:**<br>Pictures, videos, websites<br><br>**Interactive TV:**<br>Videos (longer version) | **Online:**<br>Pictures, videos, maps<br><br>**Interactive TV:**<br>Videos |
| Browsing facets | **What mode:**<br>Wikipedia entry on UN Peace troops<br>Flickr photo's<br><br>**Where mode:**<br>Other conflicts where peacekeepers<br>are active | **What mode:**<br>Other flags<br>Other emblems<br><br>**Where mode:**<br>Maps of France<br>Documentary of Paris |

**Figure 7: Example for the documentary scenario**

This documentary scenario is especially interesting from the video content annotation viewpoint. A rich information environment is needed containing the many relationships in a cultural heritage setting. This will allow users to relate the basic annotations to the many different aspects this annotation may have in different cultural contexts. The content of an artistic documentary may be described in many different ways. Intelligent search is needed to match user queries which typically come from a specific context with these many different views and aspects. For this, semantic technologies such as ontologies are particularly useful, and existing ontologies such as ULAN (artists) can be integrated into the system. The

LinkedTV user interface will support the user to search and browse the rich information associated to cultural content.

### 1.1.2.3 Scenario 3: Media arts

The media arts scenario is based at the *NUMEDIART Institute for New Media Art Technology*[1] recently founded and hosted at the University of Mons. The institute gives a space to media artists to explore different aspects of media including ubiquity and interactivity. LinkedTV technology will empower media artists to explore the interweaving of television and the Web in new ways.

The scenario introduces the artist as a mediator between a public and available multimedia content. Availability of tools for hyperlinking and annotating various forms of content, and the content itself, will constitute tremendous assets in increasing productivity, in providing a platform for experimenting novel approaches in storytelling, or more generally in stimulating creativity. In a context of reuse and repurposing, an artist can browse content in order to rapidly find the data he needs in order to create new content, or even an artistic performance. Indeed, several experiments of live content browsing seen as an art or entertainment form can be found, including the art of skilled live music remixing. Recently, richer content is experimented live by artists, and content browsers are part of the performance, facilitating retrieval, or supporting the storytelling, in settings where the screen is part of a hybrid stage.

Public feedback is an important facet here. In a world of networked creation and distribution, the artist can also benefit from live feedback of explicit (specific gestures) or implicit (individual emotions, social interactions) behavior of the public. The content can be distributed on hybrid internet-broadast services to classical setups in a home, or in more or less small cinema rooms where a more important crowd can watch and the social interaction can be more complex. Depending on this feedback, the artist can decide to modify his performance or the provided content, the public hence impacting on the interactive story that the artist mediator conveys.

### *1.2 Project objectives*

The Television Linked To The Web (LinkedTV) IP proposal aims at an innovative, integrative, and practical approach towards Networked Media which will be tackled via a set of tightly interrelated key scientific and technological research goals and measurable project objectives:

- The LinkedTV methodology aims at a comprehensive, integrative, and practical approach for annotation in audio-visual material. We will exploit and integrate state-of-the-art methods and tools (including those from other EU projects) and extend them in order to provide a fine-grained annotation not just of the whole video but also of its content on a finer basis.
- We will use external information for these annotations from many different sources related to the audio-video material which
  - was generated during the content production process,
  - which comes from external sources like Wikipedia or Google Maps,
  - from the audio track,
  - user annotations in a social network, etc.
- We will develop appropriate multimedia description schemes for annotating audio-video material with the necessary precision and granularity for LinkedTV services.

---

[1] http://www.numediart.org/

- The LinkedTV methodology and platform support the Networked Media production process in a novel way. The many currently unrelated pieces of information will be brought together for annotation purposes. On the other way, the availability of annotated audio-video information will significantly support the production of new annotated audio-video content.
- The annotations provided to audio-video content will also allow us to interlink this content with other related information of any kind (audio-video, graphics, text, data) on the Networked Media Web. We refer to this new Web-based layer of interlinked media content as the *Linked Media Layer*.
- Search engines will be able to use this annotated information in a way much similar to the way they are used for text based information.
- We will support the end user using this novel approach of annotated audio-video content and all related features (interlinked, searchable). Appropriate browsing and information presentation concepts will be defined for Networked Media for the different kinds of users and scenarios. This will smoothly be integrated into current Web browser and TV environments.
- We will support users to manage the huge amount of interlinked multi-media information in a way tailored to their requirements and needs. User profiles, filters, and other user personalization techniques will be integrated. Context awareness will be integrated which can ensure the greatest relevance for the viewer with respect to the concepts which can be selected and the content which is presented
- The LinkedTV platform will allow us to demonstrate the feasibility of our approach. It integrates automatic and semi-automatic annotation techniques with manual techniques, with knowledge management techniques in order to use other related information for annotations, and with user interaction capabilities for user-driven annotations. It will be developed as an end-to-end platform to integrate all LinkedTV services into the existing content development chain - from content producer through broadcaster and delivery network to the user device.
- To achieve this, we will repurpose the components/standards/models developed by our partners and as part of previous research projects, extending and integrating these where needed.

Our LinkedTV methodology and our platform will be evaluated in our three representative use scenarios and in related user trials.

Commercialisation of results will be supported by business modelling:

- Analyzing market and user needs and requirements regarding connected media and to extract market and user-centered use cases
- Analyzing and developing potential business models for innovative connected media applications including in particular analysis of emerging players and value chains as well as potential for commercialization through innovative connected advertising concepts.

Through industry participation, LinkedTV will transfer its results to the Web, broadcast and telecommunication industries to enable new and innovative IP-based services which can create new market opportunities for SMEs and protect and improve the competitiveness of large traditional market players in an increasingly fragmented future media landscape.

Hence, the LinkedTV project will enable a new approach for hypervideo with dynamic and personalized content enrichment on any online media device, hence making Television linked to the Web **the** ubiquitous and intuitive information delivery means of choice in the future digital society. We expect LinkedTV results to be re-applicable in future networked media contexts, such as mobile, 3D Internet or virtual worlds. The basic approach will have a lot in common with the needs in these media paradigms and the LinkedTV platform will be designed to be neutral to the chosen content delivery network as well as adaptable to the chosen end device (Web, TV, mobile etc.).

### 1.2.1 Objectives by workpackage

In the table below, we illustrate the place of each objective in our work plan, and the means for its concrete and measurable evaluation.

**Table 3: Objectives by workpackage**

| Objective | WP | Achievement | Evaluation |
|---|---|---|---|
| Apply and improve media analysis approaches for the semantic analysis and linking of audiovisual information | 1 | (1) Automatic and semi-automatic decomposition, representation, similarity evaluation and semantic annotation of video<br><br>(2) Analysis of complementary text / audio information, and use of results from different modalities for optimizing other annotation processes<br><br>(3) Labeling of content segments with event labels, identification of specific instances of objects and events, and use of these results for the semi-automatic linking of content with minimal effort | Ground-truth data setup, manual or semi-automatic annotation of available content, design of evaluation metrics and methodologies, performance of experiments to measure the effectiveness of the developed technologies, participation to relevant international benchmarking activities |
| Enable Web-scale hypervideo through the conceptual annotation of media | 2 | (1) A URI-based mechanism for addressing fragments of media content<br>(2) Lightweight annotation schemes interoperable with most of the multimedia metadata standards<br>(3) Novel information extraction techniques and disambiguation algorithms for web content | 1) Client-side and server-side implementation of media fragments that conform to the Test Cases developed in the W3C Media Fragments WG<br>2) Lightweight ontology model and API for reading and writing metadata, compatible with existing standards and conform to the Test Cases developed in the W3C Media Annotations WG<br>3) Ground-truth data setup, benchmark and evaluation campaign comparing existing information extraction tools and assessing the added value of novel disambiguation algorithms |
| Develop intuitive interfaces for Web | 3 | (1) Understanding of user tasks for accessing linked | (1) Sufficient functional and user interface requirements to |

| | | | |
|---|---|---|---|
| and IPTV-based concept browsing and content presentation in video | | information from networked, dynamic media.<br>(2) Creation of interactive presentation engine for supporting tasks in a concept-based navigation environment.<br>(3) Development of design guidelines for user interaction with linked information in a media network. | allow design of supporting software in the context of the project scenarios.<br>(2) User environment satisfies identified requirements and can be incorporated into the overall LinkedTV environment using scenario data.<br>(3) Confirmation that users' tasks are supported in the context of the scenarios through user evaluations. |
| Personalise selected concepts and content presentations to the user's profile and context | 4 | (1) Unobtrusive detection of user general interests and their semantic representation<br>(2) User behaviour pattern recognition and knowledge pulling for contextual personalization<br>(3) Enhanced state-of-the-art techniques on content-profile matching for personalized concept and content delivery | Manual or semi-automatic content and concept ratings, design of evaluation metrics and methodologies for detecting user interests and behaviour patterns, user study |
| Create an end-to-end platform for LinkedTV services | 5 | (1) Platform deployable in heterogeneous environments<br>(2) Flexible integration of all components<br>(3) Seamless cooperation of all components to generate annotated videos<br>(4) Provide clickable video playout for Tele-vision linked to the Web | (1) Evaluation of each component in the corresponding WPs 1-4<br>(2) Systematic verification of all features from end- to–end in WP5<br>(3) Evaluation from the user perspective by the trials in WP 6 |
| Realize the scenarios where LinkedTV services are used to enhance the TV/video experience | 6 | (1) An environmental / informational scenario<br>(2) A cultural heritage scenario<br>(3) A media arts scenario | 1) Internal qualitative evaluation by the partners<br>2)User trials using members of the target community<br>3) Public feedback to scenario demonstrations |
| Disseminate project results in the community | 7 | (1) Conference papers<br>(2) Journal articles<br>(3) Workshops on LinkedTV topics<br>(4) Standardisation actions at W3C etc. | 1) Number of publications<br>2) Global standing of the conferences and journals where LinkedTV results are presented<br>Attendance at and results of LinkedTV workshops<br>Contribution to standards |
| Exploit project results commercially | 8 | (1) Market surveys<br>(2) Business plans<br>(3) Exploitation plans | 1) Market reports on potential for LinkedTV services and technology<br>2) Business plans taken up in |

| | | | commercial ventures 3) Exploitation by LinkedTV partners and others |
|---|---|---|---|

## 1.2.2 Relation to call

LinkedTV focuses primarily on the target outcome a) **Digital Media Delivery Platforms.** In it, the following is stated:

*„Architectures and technologies for networking and delivery of digital media, provided through open environments enabling personalisation and high user involvement capabilities"*

LinkedTV will develop an *architecture and technology for networking and delivery of digital media* at two levels:

- for creation of LinkedTV services an end-to-end platform will provide the required components to annotate media, create interfaces to video objects, browse concepts and access multimedia presentations;

- for supporting the LinkedTV platform we will also specify guidelines and schemas for an Internet-wide metadata layer linking media to concepts and concepts to media so that video objects can be dynamically and automatically connected to collections of related Web-based media.

The LinkedTV project also partially addresses objectives of target outcome **c) Multimedia search**, which asks for:

*„Scalable, multimodal, real-time media (image, audio, and video including 3D media objects) search and retrieval technologies deployed over open platforms. Search engines that facilitate and personalize fast access to web-scale digital media objects, beyond text based indexing and retrieval capabilities of currently available search technologies."*

Our concern in LinkedTV is that future Networked Media – in which television will be a part, or all of it will be for future viewers "television" - is not yet richly interconnected, that is, it exists on a common network (generally the Internet) and hence technically media can be linked (at the data level) or associated (at the metadata level) with one another, yet these links and associations do not exist. While consuming media, and particularly complex media such as audiovisual materials, it is very meaningfully to associate fragments of the consumed media to other, conceptually related, media on the network. However, traditional search and retrieval, such as entering a text string to find media, is not intuitive in the context of watching TV programs or Web video. Rather, LinkedTV makes multimedia search and retrieval in this context much more intuitive by enabling the selection of objects in video and from this point, browsing along concept facets and accessing associated media which is packaged in an informative multimedia presentation.

The accessed multimedia content in response to the interaction with a video object and subsequent selection of a concept will be gathered from across the Web, from different sources, of different types and with different origins, with the LinkedTV personalisation components ensuring a directed filtering and selection of content and the LinkedTV presentation components ensuring a meaningful packaging of the final content in an informative presentation.

However, fundamentally LinkedTV is about enabling the next generation of audio-visual services which go beyond classical television (passive media consumption) and the current trends of on demand content and on-screen applications, and intuitively offer users **enhanced access to knowledge and on demand presentation of information in a multimedial manner within the Television experience.** This meets the expected impact of the call for:

*"Digital media/service platforms aggregators provided with innovative offers for immersive, interactive and personalised digital media."*

**We see this as the creation of a new hypervideo experience for consuming future Networked Media which will define what a future generation will know as "television",** in a similar manner to how hypertext was the instigator of the success of a particular form of accessing and browsing content on the Internet which we know as the "Web". In combination with the LinkedTV efforts in demonstration, dissemination and exploitation, and in particular the important role played in the LinkedTV consortium by the SMEs STI International, CONDAT and Noterik, this meets the expected impact of the call for:

*"Reinforced positioning of the European ICT and digital media industry, and wider market opportunities, in particular for technology-providing SME's."*

### 1.2.3 Progress beyond the state-of-the-art

LinkedTV targets the state of the art in interactive media experiences, in which Web-based approaches for PC devices are currently the most advanced. In our application to IP-based TV and Web services, we propose a system that will provide dynamic and innovative content aggregation that goes beyond any offerings that are available at present. Quite like hypertext was the key in the boom of the WWW by its one-click link-following navigation paradigm, we argue, in the application of LinkedTV functionality across networks and devices, that hypervideo will be realised with the potential to become key to the boom in the networked media domain (the Wikipedia page on Hypervideo at the time of writing contained an argument for this, see http://en.wikipedia.org/wiki/Hypervideo).

### 1.2.3.1 Web-based interactive video

The model of interaction foreseen in LinkedTV may, initially, appear to be similar to services such as Asterpix (see http://www.asterpix.com/). Asterpix, another provider of interactive video technology, recently released **Asterbot**. According to recent information supplied by the company, Asterbot automatically tags any web video with interactive hotspots on the most salient objects, which allows the user to click and acquire relevant information. The system ranks all candidate regions in the video in order of attention they receive from the camera, clusters the text around the video (title, description etc.), ranks them in order of importance and then assigns salient regions to salient topics.



**Figure 9: Asterbot allows the user to click on automatically detected salient regions and acquire information about related video clips or text in the web**

Browsing the video clips found on the official site reveals the impressive impact such a technology may have in the web community, but also demonstrates its current weaknesses. Specifically, objects of interest are often not related to the actual interest of the user and they are not efficiently tracked throughout the whole or long part of the clip. Furthermore, although the interface is designed in a rather simple fashion, it does not give a feeling of actual interaction.

Another similar and seemingly related service is Videoclix (see http://www.videoclix.tv/). There, hypervideo is used to monetarize video through clickable regions that let additional information pop up, and sponsor ads. Videoclix offers a high quality interface with more

links, more precisely set on the screen, and more things to do with them, but everything here is done manually with an authoring tool at a high cost.

blinkx BBTV (BroadBand TV - see www.blinkx.com) is heralded as another significant advance in online video. It leverages blinkx's patented technology to simultaneously deliver video over the Web and link it to the breadth of information on the Internet, adding dimension and context to the viewing experience. It uses hybrid peer-to-peer streaming and a simple point-and-click channel interface to deliver a new kind of online video: full-screen, TV-like quality and truly immersed in the Internet.

By providing a transcription of the audio stream, blinkx BBTV enables users to instantly browse or interact with online sources related to what they are watching by clicking on a word in the transcription. Current sources used include Google, Wikipedia and the Internet Movie DataBase (IMDB). However, the technology can not identify concepts which are not explicitly mentioned in the audio stream, nor handle synonyms and linguistic ambiguity (a word may have several meanings). The link to "related concepts" is basically a search on external Web sites using the chosen word as search term, leading to varying levels of relevance in the results.
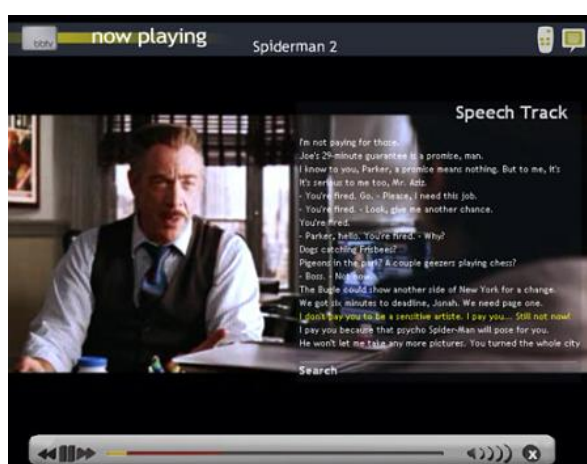


**Figure 10: BlinkxBBTV interface**

What LinkedTV aims for is a more intuitive and automated approach to enable video consumers dynamic and personalised access to associated content based on concepts in the video. The LinkedTV consortium strongly believes that the work to be carried out in the course of the project will go at least one step beyond what is available to date and allow for a much higher degree of interactivity by providing more precise and richer associations between salient objects and conceptual descriptions, and by facilitating intuitive access to related web information.

### 1.2.3.2 IP-based interactive TV

Classic IPTV has established itself, typically as part of "Triple Play" offers, as a successful means for telecommunications operators to offer new types of services around television such as EPGs, programming on demand and live TV pause.

Web-TV convergence in the IPTV market has been markedly less successful to date, with offers focusing on either forcing Web content onto a TV screen (resulting either in poor results or requiring Web authors to write new pages in TV-friendly markup) or classic linear TV for the big screen being streamed onto PC screens. The alternative approach, commonly referred to as "Interactive TV (iTV)" packaged additional content with TV programming which was produced manually in advance at a disproportionate cost. One major STB

platform, Multimedia Home Platform (MHP), required services to be developed in Java and was remarkable only in its complexity.

Currently, Web-TV is approaching a critical mass on the market, making the potential of truly using the Web while watching TV more realistic now. Digital IPTV offerings such as HbbTV in Germany or YouView in the UK indicate that in the next years, TV-based Internet connectivity and application access to data from the Web will become more commonplace.

What functionalities do current IPTV platforms provide?

Firstly, Web-based video services are being integrated into the IPTV experience, with the aim that in the future it may not be visible, nor relevant, to the viewer whether the viewed content is coming from a broadcast network or over the Internet, and whether its source is a broadcaster or a Web hoster such as YouTube or Netflix or any other kind of source. What matters most is the content's relevance to the user and his/her respective needs and requirements, plus trust in the provider that the information supplied is accurate and no misuse of user data or usage patterns occurs.

Secondly, another TV trend is towards integration of Web content onto the TV screen, mainly through widgets, which are lightweight self-contained content items that make use of open Web standards (HTML, JavaScript) and the back-channel of the STB to communicate with the Web (typically in an asynchronous manner). Yahoo and Intel, for example, presented their Widget Channel at the CES in January 2009 (see http://news.zdnet.co.uk/communications/0,1000000085,39586222,00.htm, dated 30 December 2008), where Web content such as Yahoo news and weather, or Flickr photos, could be displayed in on-screen widgets on TV. This widget platform is supported by various manufacturers under the name ConnectedTV.



**Figure 11: Yahoo!s Widget Channel with Flickr, weather and stocks widgets**

Finally, another significant trend is social TV and personalisation, with content recommendation and EPG personalisation making the TV experience more relevant to individual viewers. BBC iPlayer for example, or various complementary TV applications like Boxee, now allow for example for viewers to link their viewing activity into their "social graph" and send or receive recommendations for programs to and from their friends. Looking more to the future in this area, the Dutch project iFanzy[2], and now an EU project NoTube[3], research how to improve TV personalisation even further through using semantic metadata. These complement LinkedTV as their focus is on complete TV programs, while LinkedTV

---

[2]  http://www.ifanzy.nl
[3]  http://www.notube.eu

looks more finely grained at personalisation at the level of video objects and associated content.

None of these trends however tie Web content, as delivered by widgets or displayed on an in-TV browser, any more tightly with the content of the currently viewed TV program, due to the lack of richer annotation of the programming (beyond EPG metadata at the atomic program level). An indicator of where this could proceed in the next years can be seen with Blinkx which currently offers a PC-based download to access Web TV material. Through subtitles and speech recognition Blinkx offers viewers the ability to select concepts (as in words spoken) in the material and links out to Google search, Wikipedia articles etc. tied to those words. However, a contextual understanding of the natural language is missing (whether Paris now refers to the city or the person, and which city - in France or in Texas - in fact?) making this in many cases hit-or-miss.

LinkedTV will hence help push IPTV beyond its current boundaries by providing object-level annotation of TV programming which includes precise conceptual identification (of "Paris" or any other term). It will also develop the infrastructure to enable new IPTV services based upon this: (1) clickable video player will provide the means for viewers to select objects in-TV; (2) the Linked Media infrastructure on the Web will provide the means to collect relevant media associated to the concept selected; (3) LinkedTV front-end display layer will make possible the appropriate presentation of the collected media to the viewer.

What is currently lacking is:
1.  a framework for identifying and extracting concepts from media,
2.  defining associations between media based on those concepts, and
3.  creating intuitively accessible multimedia presentations based on those associations.

This is where LinkedTV comes in. Before we describe our approach and objectives in more detail, let us first turn to the state-of-the-art in these separate fields.

### 1.2.3.3 Multimedia analysis

State-of-the-art technologies in the field of multimedia analysis developed in the scope of current and recent European research projects, such as GLOCAL, VidiVideo, X-Media, MUSCLE, MESH, Imagination and Portivity, are promising and promote further research and new application targets. For instance, methods for single medium information extraction from images and texts exist [37][38] and cross-media mining applications are currently emerging as a result of these projects [39][40]; however, they lack the ability to use the full power of contextual information in their analysis, either as prior to improve output confidence or as a way to improve the underlying processes. Furthermore the implemented fusion processes are mostly based on combining the single-media results rather than on continuous recursive cross-media interaction. Another promising direction that very recently started to emerge is the use of not only object and scene labels but also of events, for the purpose of labeling and indexing multimedia content [41][42]. Nevertheless, the labeling and indexing of content with events, as well as the detection of specific instances of objects and events, rather than more general labels of object / event categories, are still in their infancy; significant research efforts are needed before such tasks can be reliably performed on large volumes of data belonging to broad domains.

Large scale multimedia analysis as seen in the scope of the project is related to any technology that helps to organize "archives" by their multimedia content. One problem with current approaches applied specifically to visual content is the reliance on visual similarity for judging semantic similarity. That's why long established solutions like "Google Image Search" and "Yahoo! Image Search" are based on textual metadata accompanying images rather than visual similarity. Nevertheless, beta versions of public domain search engines based on visual similarity have started to appear, like Riya that incorporates image retrieval and face recognition for searching people and products on the web. Video sharing, as

implemented by YouTube, put emphasis on the need for Content Based Image Retrieval (CBIR) on multimedia data (video, audio and text). The hypervideo notion is used to characterize tools like Asterpix, Videoclix, Hypersoap,, Klickable, Overlay.TV and Blinkx, to name but a few. All these tools produce videos that contain embedded, user clickable regions that allow navigation between video or web information. All of them but Asterpix are totally based on human authoring. Asterpix automatically provides objects of interest, but in many cases it fails to capture the essential depicted regions/objects.

Multimedia analysis in the scope of the project will tackle the main technological barriers of how to detect content segments of interest in space and in time, label them with rich semantic descriptions (including event- and object-instance-labels), and finally link them to e.g. another part of the same or some other video, or to external sources. In the framework of the project we will exploit/extend state-of-the-art technologies in single modality processing and will make a step in fusing information from diverse modalities. Specifically, we foresee contributions in the following fields:

- Effective automatic or semi-automatic decomposition of content to meaningful segments (e.g. scenes, sub-shots corresponding to elementary human or non-human actions, spatiotemporal or spatial regions of interest). It will use current segmentation techniques that require minor user intervention and adapt them to the annotation needs of the LinkedTV annotation tool(s).

- Representation of content segments using extensions of SoA low- and mid-level descriptors (e.g. sub-shot activity areas, region- or interest-point-trajectories over time, space-time pyramids of features, maps of facial regions), and definition of fast similarity-based matching techniques that will exploit the developed content representations for identifying identical or similar instances of segments.

- Computationally efficient clustering, learning, and dimensionality reduction techniques. These technologies will enable the annotation of visual content in broad domains (as opposed to the state-of-the-art, which is mostly limited to narrow domains), i.e. the automatic representation of each content segment with a model vector in an appropriate high-dimensional semantic space (a space defined by a multitude of object and scene labels), and the recognition of faces.

- Audio processing as a tool for automatic segmentation of the audiovisual content based on information in the audio, such as speech/non-speech detection and speaker segmentation.

- Speech recognition as a tool for generating textual annotations of the audiovisual content for both enrichment of manually generated metadata and as a prior for other multimedia analysis tools. For example, the mentioning of an entity in the speech provides some evidence that an image of the entity may be present; the mentioning of a person can be used as a prior for speaker identification. Also speech transcripts are used as an information source for keyword/concept suggestion for human annotators early or later in the loop.

- Speaker identification based on priors from multimedia annotations. Using evidence from available annotations such as persons automatically detected or manually annotated, persons mentioned in speech recognition transcripts and speaker segmentation, speaker models based on effective cross-media interaction can automatically be generated incrementally.

- The combined use of all the above analysis results for eventually achieving an event- and instance-based labeling of visual information, i.e. for annotating video with event labels and for identifying specific instances of objects and events, in contrast to the

state-of-the-art, which is mostly limited to detecting generic object- and scene-category labels.

- Novel interfaces for browsing and retrieving information related to the media, exploiting the results of the aforementioned analysis approaches. These will be part of an annotation tool for hypervideo, which will support users in the semi-automatic linking of content with minimal effort.

Exploitation of existing methods and new ideas in the above fields will eventually bring us one step beyond current frameworks and enable a novel and complete video representation. For example, in the case of the "Jeugdjournaal" story, the LinkedTV platform will track information about what occurs in the shot (interesting regions and events), from whom (persons' names), when (timestamp) and where (spatially or temporally). Detected information will be linked to other events in the same sequence (e.g. a past or future scene), to other modalities (e.g. the news story on a website) or to specific repositories (e.g. cultural knowledge about the event occurring in the shot) and thus giving the user the opportunity to retrieve diverse information or browse the content in a context-sensitive way.

Quite to the contrary of current hypervideo implementations, that annotate content only manually, LinkedTV will provide semi- or/and fully- automated annotation tools as well as intelligent interaction with both internal (e.g. segmented and identified objects of the video) and external video objects (e.g. external hyperlinks to user-selected boxes). In addition, multimedia analysis exploiting diverse media will also aid in bridging the gap between traditional visual analysis methods (i.e. person/face detection, tracking, etc) and the fact that they do not take into account the users' context and perception or any kind of social information.

All in all, LinkedTV will enable visual similarity and contextual relations methods to play the key role in providing successful results and go beyond typical tag- or text-based approaches (e.g. YouTube, Blinkx). Its automatic face/object detection, identification and tracking features together with the interactive segmentation and tracking functionalities will make it easy to annotate content on-the-fly and go beyond typical approaches (e.g. Videoclix, Overlay.tv, Asterpix). Finally, the introduced hypervideo-based mode of interaction will allow interactive link-following navigation in the Networked Media domain and will aid the mainstream trend of producing easily searchable video content.

## 1.2.3.4 Web data mining

Web mining, particularly Web content mining, focusing on identification and extraction of information from the textual content of a web page, is a popular research area of the last decade. There is a wide variety of different approaches ranging from simple wrappers mapping extracted information on the pre-defined data structures to complex extraction tools employing advanced machine learning techniques and deep linguistic processing [6].

These approaches however exclusively rely on the information provided by the textual content of the document. Apart from Web content mining there are areas dealing with web site structure, mostly exploiting the hyperlink structure using graph-based methods, and web usage mining that analyzes server logs and tracks user behaviour on the web site. The idea of Web content mining in connection with multimedia content analysis is largely a novel research direction. It has been experimented with in FP6 projects such as K-Space [7] or Boemie [8]. Current approaches typically focus either on web documents with pre-defined structure (e.g. online web reports for football matches) or perform simple web searches for unstructured content - whole relevant documents (without fine-grained analysis provided by information extraction methods). LinkedTV will go one step beyond the state-of-the art in Web mining for multimedia by **making mining results available in the form of a semantic annotation**.

Apart from Web content mining LinkedTV will utilize **Web usage mining** that analyzes server logs and tracks user behaviour on the Web site. Part of the interaction with the LinkedTV platform will be subject to Web usage mining contributing to the personalization experience offered by the platform.

In a wider sense, web mining can also involve additional information linked to the web infrastructure, other than text, hyperlinks and user behaviour data. In particular, distributed structured information available over the web infrastructure, such as *Linked Data*.

Information mined from the Web can be utilized in LinkedTV in two processes – in the multimedia annotation process and during the content presentation.

In the **multimedia annotation process** the identification of concepts and objects based on the low-level audio-visual analysis can be improved by drawing in additional information extracted from the Web. Web content mining can provide a complementary source of evidence for the low-level audio visual analysis by helping to identify context and hence allow for disambiguation - e.g. by discarding some concepts identified by the audio-visual analysis and boosting the importance of others.

In **content presentation**, new information from both unstructured and structured content will be collected and filtered so that the viewer can be presented information relevant to the current storyline of a shot, scene or the whole broadcast. LinkedTV's work in this field will be divided into two main streams: one focusing on filling given templates with information found on the Web, and the other gathering diverse information related to the broadcast or the programme. Novel mining techniques will allow to create knowledge bases for the given topic area and to decide on the fly among different (push or pull) information delivery models based on the nature of the broadcast.

**Technologically**, Web scale data mining poses a challenge not only from the point of the development of scalable algorithms, but also from the point of acquiring the data (crawling).

In this respect, the consortium will consider two main options for data acquisition. The first option builds upon existing cloud-based services for Web crawling. For example, the Web crawling service offered by 80legs requires no upfront investment costs and the price of $2 per million pages crawled and $0.03 per CPU-hr used[4]. This approach allows us to data mine up to one billion Web pages and is cost-effective compared to the consortium acquiring its own computing and connectivity resources for the trials. The second option considers obtaining most data from the Linked Data Cloud and using primary textual documents only in a limited way. In this scenario, the Linked Open Data (LOD) cloud will serve as the prominent source for feature extraction algorithms with text-mining playing a complementary role. One of the key resources to consider are DBpedia for LOD efforts and Wikipedia for text mining. Both options will be assessed and the most viable one implemented.

### 1.2.3.5 Multimedia search and retrieval

Retrieval of multimedia objects is generally carried out in one of two ways:

- "content-based"
- "text-based"

The content-based retrieval uses the analysis of low level features [9] to represent the multimedia content. Such features are usually based on colour distribution, texture, edge, motion, geometry etc. to describe the visual content. The indexing structure of low level features allows retrieval employing the "query by example" methodology (i.e. finding pictures of an object based on an example photo of that object). More complicated approaches include segmentation of the example image so the user can query a system by using as input only a specific region (i.e. the object of interest) [10]. In the field of segment-based image

---

[4] http://www.80legs.com/

annotation, several research approaches adopt part-based visual features, either grids ([11]) or segmented regions ([12]), but they do not aim to establish the correspondences between individual parts and concept labels.

On the other hand, text-based retrieval uses the indexing of media according to text that can be associated to it, such as titles or descriptions in associated metadata files, or text found close to the media on a Web page. Such retrieval methods are also exploiting external resources to define concepts with the usage of synonyms, hyponyms and textual hierarchies. Related work is oriented towards classifying extracted low-level features to a set of visual words, like in [13], where visual categorization is achieved using a bag-of-keypoints approach, or in [14], where the classification initially relies on the extraction of Local Interest Points based on the extraction of SIFT features, quantized using a visual dictionary. Other related approaches employ ontologies to translate the textual information into a concept-based meaningful representation, like in [15], where a region-based approach using MPEG-7 visual features and ontological knowledge is presented. Although text-based retrieval can be considered as more reliable, it is highly based on the quality of the annotations, which are usually noisy (e.g. tags from different users), or they do not describe adequately the visual content.

More recent works [16] combine the aforementioned techniques by employing fusion methods of the heterogeneous information and promising results have been presented. One generic approach to multimodal concept detection is combining multiple single-feature or single-modal classifiers by ensemble fusion. Each individual classifier uses statistical models like a GMM or SVM and operates on a single feature or a small set of features. The detection scores from individual classifiers are then fused using a linear mixing function or a discriminative classifier (i.e., SVM) [17]. Different approaches for image-level annotation in [18] have been proposed, by firstly establishing correspondences between salient (or representative) regions and concept labels. However, despite the recent advances in multimedia indexing, the conceptual representation of the content still remains inadequate for sufficiently exact retrieval purposes.

To achieve a more reliable concept based representation, the so-called "semantic gap" between low level feature descriptions and high level conceptual association must be filled through a multimedia annotation. Consequently, multimedia annotation could be an additional invaluable information, which can be combined with the already available textual and visual metadata in order to achieve more meaningful multimedia indexing and concept retrieval, bridging in this way, the semantic gap.

*Multimedia Annotation*

The recent efforts towards knowledge-based multimedia analysis and interpretation are characterised by the use of ontologies, which provide the means to support explicit and machine processable semantics.

Well known efforts on multimedia ontologies include those related to the MPEG-7 standard, which constitute the main standardisation effort with respect to multimedia content descriptions. MPEG-7 can be used to create complex and comprehensive metadata descriptions of multimedia content. Since MPEG-7 is defined in terms of an XML schema, the semantics of its elements have no formal grounding. In addition, certain features can be described in multiple ways. In order to make MPEG-7 interoperable with domain-specific ontologies, the semantics of the MPEG-7 descriptors also need to be expressed formally in an ontology.

For alleviating the lack of formal semantics in MPEG-7, four multimedia ontologies represented in OWL and covering the whole standard have been proposed. Chronologically, the first initiative was taken by Hunter who proposed an initial manual translation of MPEG-7 into RDFS [19]. In 2004, Tsinaraki et al. have proposed the DS-MIRF ontology that fully captures in OWL DL the semantics of the MPEG-7 MDS and the Classification Schemes [20]. In 2005, Garcia and Celma have presented the Rhizomik approach that consists in

mapping XML Schema constructs to OWL constructs following a generic XML Schema to OWL together with an XML to RDF conversion [21]. In 2007, Arndt et al. have proposed COMM, the Core Ontology of MultiMedia for annotation [22]. These four ontologies have also been compared [23].

Another issue is how to make use of such explicit semantics, connecting low level content descriptors with high level domain concepts. For example, in the aceMedia [24] project, a Visual Descriptors Ontology is developed to cover MPEG-7 visual descriptors, which can then be linked to domain specific ontologies for the purpose of assisting semantic analysis and retrieval of multimedia content. In general, Thonnat et. al [25] show that an ontology can be used to guide and assist semantic image analysis by capturing three types of knowledge: 1) domain knowledge (concepts and relations of importance), 2) anchoring knowledge (mapping symbolic representations to visual data) and (3) knowledge related to image processing (which algorithm to apply, which parameters to use, etc.). It is expected that such classification can be applied in other kinds of multimedia analysis.

From the industrial side, the Metadata Working Group (MWG) was formed in 2006 as a consortium of leading companies in the digital media industry with the intent of publishing technical specifications that describe how to effectively store metadata into digital media files [26]. Finally, the W3C Media Annotations Working Group [27] aims at developing a simple lingua-franca between most of the multimedia standards.

We will tie the multimedia annotations to a growing body of concept-centred metadata on the Web known as "Linked Open Data"[5]. This is an initiative to create a Web-scale layer of metadata about concepts using the principles of the Semantic Web to:

- give every concept an URI
- describe concepts using this URI and shared ontologies which provide the definition of used classes and properties
- link descriptions together by shared concepts (e.g. the object of a statement about one concept is the subject of another statement on the Web) and by linking URIs (by stating that two different URIs refer to the same concept)

The effect of Linked Data is the ability to access machine-processable information about concepts on the Web using Semantic Web technologies (RDF, SPARQL) just as one can access human-readable information about things in a Web browser using the Web technologies (HTML, HTTP, URLs). By tying multimedia annotations to Linked Data concepts we will enable a *Linked Media Layer* where a machine can identify the related concepts to a particular media segment and from this, allow for user browsing along concepts and eventually the retrieval of related media to the browsed-to concept, reusing the existing and growing Linked Open Data on the Web (as well as supporting its creation in our own work).

### Multimedia Fragment Addressing

Providing a standardized way to localize spatial and temporal sub-parts of any non-textual media content has been recognized as urgently needed to make video a first class citizen on the Web.

Related work can be traced back to hypermedia research. An hypermedia document [28] refers to a collection of information units including information about synchronization between these units and about references between them. Temporal and spatial dimensions are typically included, whereas references can be made between parts in both dimensions. The issue of linking in hypermedia is discussed in [29,30,31]. Linking within multimedia presentations, within and among linear and non-linear multimedia presentations is discussed

---

[5] http://linkeddata.org

in [31]. [29] discusses links in time-based presentations and proposes a method to specify the context of links, i.e., what happens with the source or destination presentation of a link when it is traversed.

Hypermedia presentations consist of both static and dynamic media objects which are grouped together in so-called composite entities. Parts of these entities, identified via anchors that provide hooks for links, can be linked with each other and the behaviour of source and destination entities can be defined (e.g. shall the source video be paused or replaced). The ideas discussed above were implemented in the Synchronized Internet Markup Language (SMIL) [32], a W3C recommendation that enables the integration of independent multimedia objects such as text, audio, graphics or videos into a synchronized multimedia presentation. Within this presentation, an author can specify the temporal coordination of the playback, the layout of the presentation and hyperlinks for single multimedia components. The latest version of SMIL provides a rich MetaInformation module which allows the description of all elements of a SMIL document using RDF.

Previous attempts for addressing multimedia fragments include also non-URI based mechanisms. For images, one can use either MPEG-7 or SVG snippet code to define the bounding box coordinates of specific regions. Assuming a simple multimedia ontology available (designated with "mm" the following listing provides a semantic annotation of a region within an image:

```
<http://example.org/myRegion>
     foaf:depicts http://dbpedia.org/resource/Eiffel_Tower;
     rdf:type mm:ImageFragment;
     mm:topX "40px";
     mm:topY "10px";
     mm:width "100px";
     mm:height "100px";
     mm:hasSource <http://example.org/paris.jpg> .
```

However, the identification and the description of the region is intertwined and one needs to parse and understand the multimedia ontology in order to access the multimedia fragment.

URI-based mechanisms for addressing media fragments have also been proposed. MPEG-21 specifies a normative syntax to be used in URIs for addressing parts of any resource but whose media type is restricted to MPEG. The temporalURI RFC [33] defines fragment of multimedia resources using the HTTP query parameter thus creating a new resource. YouTube launched a first facility to annotate parts of videos spatio-temporally and to link to particular time points in videos [34]. It uses the HTTP URI fragment delimiter but the whole resource is still sent to the user agent that just perform a seek in the media file. In contrast, we will research and implement a solution where only the bytes corresponding to media fragments will be sent over the network while being still cacheable.

In LinkedTV we will pursue the work of a standardised media fragment addressing scheme which will be supported in the LinkedTV platform, as part of the W3C Media Fragments Working Group[6].

## 1.2.3.6 Personalisation

Current personalisation efforts focus on personalisation of the audiovisual material presented to the user (e.g. TV content recommendation by EPG, profile-based ranking of Web video results) or of the programming itself (in terms of selected temporal video segments, e.g. to

---

[6] http://www.w3.org/2008/01/media-fragments-wg.html

filter a news broadcast down to the stories of interest to the viewer). In LinkedTV, the focus is rather:

> 1. to filter the objects browsable within a TV program and the concepts they are associated to according to viewer's interests and viewing context;

> 2. to select the relevant content for that concept and adapt its presentation according to the viewer's interests and viewing context.

## *User profiling*

Related efforts in user profiling of multimedia content is limited to low-level profile representation, identifying preferences mostly in video genres and programme schedule. However, every user has a unique preference background, specialized on each area of interest, interchanging through time [34, 36]. Furthermore, preferences are not binary, since items of preference carry a distinct weight of participation to the user preference space. Breaking down area-specific user preferences would require frequent item set generation techniques that have been proposed for handling real-valued features for transactions that are based on quantizing the feature space [35].

The LinkedTV framework aims at identifying and handling more sophisticated user preferences based on advanced multimedia analysis and enriched metadata mining from Web resources. These preferences should be captured unobtrusively and be stored in lightweight semantic structures to provide computationally efficient, semantic content filtering for the PC, TV or handset. In addition, an efficient scheme of preference weighting, which is updated in time, is important, in order to distinguish the most prominent preferences in context and discern from long and short term interests. While lightweight storage and efficient handling of semantic user profiles in resource-challenged environments has been addressed [43][44], LinkedTVs contribution on user profiling would extend this work to the following fields:

- Unobtrusively tracking user transactions and combining them with mined Web domain information to extract enhanced metadata.
- Semantically classify and represent user preferences with a standard semantic formalism.
- Implicitly inferring complex user interests by means of frequent preferences recognition.
- Enhance known profile matching techniques in order to produce fast and server-independent filtering and ranking of delivered concepts and content by means of fuzzy semantic reasoning.

## *Contextual adaptation*

Personalized content retrieval takes into account initially the collective set of extracted and semantically represented user preferences. However, managing the general user preferences to the content delivery process would hinder filtering performance and produce obfuscating and out of context results. Advanced personalization requires that out-of-context preferences would be disregarded in order for content delivery to produce specific context-aware recommendations [35]. Adapting user preferences and recommending content within context in a multimedia environment is currently oriented towards contextualizing the time, space, task and state of mood of the user, in aspects that are related to the recommended services.

The research conducted within LinkedTV will focus on dynamically updating the context of ongoing search and retrieval actions, with respect to pulling subsets of the user?s long-term interests and of the domain knowledge available through the exploitation of extracted content metadata in order to adapt the proposed concepts and content to the user?s context. Refraining from using large and complex knowledge bases to outline the physical, virtual and logical

data of the users [45], LinkedTV will rather focus on identifying and pulling only the logical data that are in alignment with the physical state and virtual information of the user, with an interest in keeping the contextualization process light and agile. More specifically, the LinkedTV framework expects contributions in the following fields:

- Recognizing the user behavioral patterns among different areas of interest in the user profile, in order to infer persistent semantic relations between knowledge areas that would indicate context-adapted user interests, taking into account both content as well as time and location context.
- Enhance knowledge evolution and pulling techniques to take into account trends and topic patterns of the context based on the rich metadata available in the Web 2.0 and Semantic Web.
- Produce contextualized personal content delivery as an intersection of contextual domain knowledge and context-aware user profiles.

*Behavioural tracking*

User behaviour, including in small collective settings (family), and focus of attention tracking will be introduced as additional facet enriching the media annotation, as well as the user individual or collective profiles, beyond explicit preferences, or preferences based on usage logs put in context.

Detecting behavioural patterns indicating strong expressed emotions (f.i. laughter) will be proposed. Behavioural patterns involving several people will also be considered, one role of media being to generate interaction between people, such as educational programs prompting questions by children to their parents, of political debates eliciting lively discussions among friends. Several approaches are available in the signal processing artillery, including computational models of human attention that will allow estimation of the rarity of tracked behavioural features. All these will allow the research on joint models of user reaction and attentive behaviour, in relation with both the low-level features of the media file frames and objects, and the extracted high level concepts related to these objects or the media content in general. Finer grained user behaviour capture and analysis will allow the extension of preference collection and hence user profiles down to the level of video frames and objects, as well as taking into account the semantic of behavioural responses to the media. Besides, behavioural reactions will also participate to the media annotation at the semantic level, as they provide information on the emotional reaction that the content elicits. This research will make use of state-of-the-art capture devices, such as depth cameras for instance (f.i. the Microsoft Kinect). In a perspective of natural interaction, such devices may be much more convenient for explicit interaction like manual annotations of frames or objects within the frames, as well as implicit interaction as suggested above.

## 1.2.3.7 User interface

LinkedTV aims to develop an intuitive user interface, which can be used to access semantically related information of audiovisual content. This requires the integration of video content with interactive graphics. This topic has been widely studied in interactive television. As a result, digital television allows integration of the interactive content with the live video. For example, the European digital television standard Multimedia Home Platform (MHP) allows mixing of background, video, and interactive applications using so called On Screen Display (OSD). This allows the development of enhanced TV programs, which combine audiovisual content with related information.

Current web video services are primarily targeted for desktop PC computers. The web sites rely on mouse-based navigation. The video is shown by the video player either as a separate part of the user interface or in full screen mode. In the former case, synchronization of the

user interface with the video content is difficult, while in the latter case only limited additional information can be displayed. Needless to say, these types of user interfaces are unsuitable for non-desktop devices, such as television.

In terms of video playout, **ClickVideo** technology is an example of the cutting edge of the state-of-the-art in this area. It makes it possible to identify parts of the video which become clickable objects in a video stream. These new objects can be addressed for the purposes of annotation or user-defined interaction. Also, they can be used in an automated method by mapping ClickVideo to any XML component which could identify relevant content. The created hotspots become clickable video objects that can be aligned with other media, such as audio, video, images or text. Extracted hotspots can be used to create video hyperlinks to relevant content, or a set of hotspots could be defined by the user for the purpose of bookmarking fields of interest, thus giving a user the possibility to define live- or on-demand video-bookmarks.

Tracking hotspots is a vital functionality of the ClickVideo component. In this way "hotspotted" objects in the video sequence will be tracked according to a set of parameters. Thus important video sequences can be identified and mixed to provide personalised sets of video bookmarks.
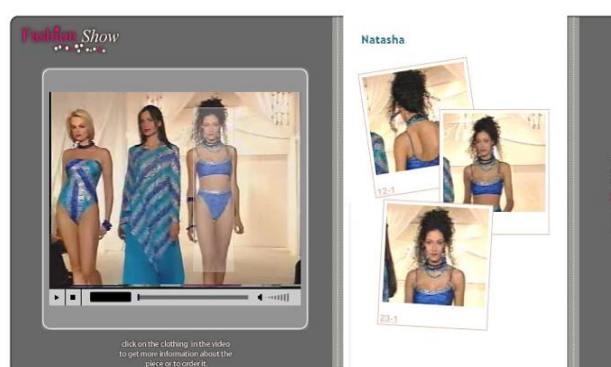


**Figure 12: example of ClickVideo interface**

In the screenshot provided above, a use case of ClickVideo technology is shown. The video shows models on a catwalk during a fashion show. A human editor has provided hotspots based on one of the models in a specific video-sequence. Once highlighted by the user, relevant close-ups of the specific selected hotspot are shown outside the actual video. Within the scope of LinkedTV, the user interface will be able to have identified hotspots in the display framework itself, rather than to align relevant objects outside a video-player.

In particular, in LinkedTV we will provide a MediaCanvas API and support playout of clickable video. The Media Canvas API will provide a layered structure, whereby any available element can be retrieved on the Canvas itself such as a subtitling element, a logo element, or any other XML driven component. LinkedTV will use this technology innovatively to provide viewers with the interface to find and select on-screen objects while viewing video, browse concepts and access multimedia presentations, in-screen.

### 1.2.3.8 Multimedia presentation

Generally, the task of assembling different media resources into a meaningful, synchronized presentation today is a manual one, undertaken by other professionals. The benefit of at least partially automating this process is clear, in terms of time and hence money saved. Within the semantic multimedia research community, there has also been consideration of the use of the semantic annotations applied to media in determining its presentation with respect to other media. Additionally, some work has been done also on the (semantic) description of design rules and templates to facilitate the "intelligent multimedia presentation" process.

Work on so-called Intelligent Multimedia Presentation Systems (IMMPS) dates back to the 1980s, and was formalised by Bordeghoni et al in a Standard Reference Model [38]. Many later systems took care to compare themselves to this model. While it defined the need for formal knowledge in the multimedia generation process, there had been no agreement on which model to use. With the emergence of Semantic Web technologies, RDF and OWL were used in development of such systems, while their integration into the process was generally quite restricted. Furthermore, they were not expressive enough for the design rules and constraints which needed to be expressed. An exemplary research work in this field is Cuypers, by consortium member CWI, which models the multimedia generation process in five phases, allowing for backtracking from one phase back to the previous one [39]. Rhetorical Structure Theory (RST) is used to determine effective communication of media in a presentation and ontologies have been introduced to formalize the domain and design knowledge applied in its generation [40].

This work will be extended, specialized in the topics covered by the LinkedTV scenarios and applied to the context of IPTV in the project.

## 1.2.4. Comparison with other projects

Regarding the state of the art analysis, it is also important to be aware of and have links to the other recently finished and ongoing projects in the fields of research and development overlapping or complementing the intended work in LinkedTV. This will ensure that there is no duplication of effort, that existing results can be applied or extended, and that parallel research activities may be brought together in co-operations for mutual benefit. Here LinkedTV has a clear advantage in that many of its consortium members are directly involved in the relevant projects. Here we mention only a few key projects, a longer list is provided as an Appendix at the end of the proposal.

**Table 3: Key projects compared to LinkedTV**

| Project description | Similarities and differences with LinkedTV |
|---|---|
| **GLOCAL** (FP7-CP) - Event-based Retrieval of Networked Media. GLOCAL offers a Semantic Web based platform for multi-media retrieval which uses events as the primary means to organize and index media, e.g., photos, videos, journal articles. A local dimension enables the mapping of tags (conceptualizations) to media (personal experiences), while a global dimension enables the sharing of event descriptions thus enabling social sharing and networking of events, across similar events, thus providing a common way to index media. | GLOCAL also applies conceptualization to index and retrieve multi-media materials, but 1) is organized on the basis of events and 2) videos remain again weakly integrated as already explained in the introduction of the LinkedTV proposal. Objects and scenes in videos can not be identified, connected to context and retrieved. LinkedTV will provide this missing gaps to identify objects in videos, link them to different types of context, retrieve them and demonstrate their use in a hypermedia player. As LinkedTV deploys existing standards such as HTML5, MPEG7, MPEG21 and SMIL it would be well suited to enhance platforms such as GLOCAL for video materials. |
| **ALICANTE** (FP7-CP) — Media Ecosystem Deployment Through Ubiquitous Content-Aware Network Environments. Alicante allows users to access multimedia services in various contexts, and  to share and deliver his | ALICANTE also uses meta data, context and user preferences to retrieve all kinds of multi-media materials, but videos remain again weakly integrated. LinkedTV provides all functions for the end-to-end process to retrieve |

| | |
|---|---|
| own audiovisual content dynamically, seamlessly, and transparently to other users. This is realised by providing an open, metadata-driven, interoperable middleware for the adaptation of distributed media resources to the user's preferences and heterogeneous contexts enabling an increased quality of experience. | concepts, to weave materials with other resources from the Web and a hypermedia player to allow new user experiences. The objects are identified not only on a frame level, but also for scenes and video sequences. As LinkedTV deploys existing standards such as HTML5, MPEG7, MPEG21 and SMIL it would allow to enhance platforms such as ALICANTE for video materials. |
| **K-Space** (FP6-NoE) — K-Space focuses on creating tools and methodologies for low-level signal processing, object segmentation, audio processing, text analysis, and audiovisual content structuring and description. It builds a multimedia ontology infrastructure, analyses and enhances knowledge acquisition from multimedia content, knowledge-assisted multimedia analysis, context-based multimedia mining and intelligent exploitation of user relevance feedback. In this process it creates a knowledge representation for multimedia, distributed semantic management of multimedia data, semantics-based interaction with multimedia and multimodal media analysis. http://www.k-space.eu/ | LinkedTV will build on the results from this NoE, since four partners (CERTH, CWI, Eurecom and UEP) already collaborated here. LinkedTV has a wider goal than just the focus on media analysis. K-Space did not have specific use cases while LinkedTV will further develop analysis and annotation technologies tailored for the needs of Connected Media Experiences and    evaluated in three scenarios. |
| **aceMedia** (FP6-IP) — The main technological objectives of aceMedia were to discover and exploit knowledge inherent to the content in order to make content more relevant to the user; to automate annotation at all levels; and to add functionality to ease content creation, transmission, search, access, consumption and re-use. In addition, available user and terminal profiles, the extracted semantic content descriptions and advanced mining methods were used to provide user and network adaptive transmission and terminal optimised rendering. http://www.acemedia.org/aceMedia | LinkedTV will build on the main outcome of this project, the aceMedia toolbox, while extending the technologies to fit with the needs of Connected Media Experiences and three specific scenarios. In particular, LinkedTV focuses on finer grained annotations of media fragments with individual concepts and on the integration of Web-based related media. |
| **LIVE (FP6, CP)** — Staging of Media Events. The LIVE project seeks to change the linear approach to live sports broadcasting by delivering digital technologies and interactive content – through personalised on-demand sub-channels. | In LIVE the focus was solely on a live TV production of an inter-linked multi-channel TV bouquet. The channels were linked by hand of professional TV producers according to the flow of the sport events. For the manual (online) and semi-automatic (offline) annotation of broadcast content (live streams / archived clips) two dedicated tools were developed, which might be re-used and adapted for usage in LinkedTV. |
| **PrestoPRIME** develops practical solutions | Related to LinkedTV, PrestoPRIME will |

| | |
|---|---|
| for the long-term preservation of digital media objects, and find ways to increase access by integrating the media archive with European on-line digital libraries in a digital preservation framework.<br>The project aims to establish metadata interoperability between audiovisual archives, cultural heritage institutions, the Semantic Web and content portals, with services for metadata conversion and deployment. | create a metadata conversion and deployment toolkit, with a novel and efficient process for metadata vocabulary alignment. PrestoPRIME focuses on aligning manually created vocabularies that in many cases is limited to programme level descriptions. The Data Mining activities of LinkedTV (see 1.2.2) will go much further than processing these manually added labels. In connection with semantic concepts and entities extracted from the low-level audio-visual analysis, information from various outside sources will be collected and filtered. The user will be presented information relevant to the current spatio-temporal segment of a frame, which represents a much finer granularity. |
| **CHORUS & CHORUS**+ — CHORUS was a Coordination Action which aimed at creating the conditions of mutual information and cross fertilisation between the projects that run under Strategic objective 2.6.3 (Advanced search technologies for digital audio-visual content) and beyond the IST initiative. CHORUS+ is the continuation of CHORUS in FP7. http://www.ist-chorus.org/ | Partners of CHORUS and CHORUS+ (CERTH) participate in LinkedTV and therefore the project will benefit from the useful experience and the results of CHORUS / CHORUS+. More specifically, LinkedTV will reuse the state of the art reports on audio-visual search technologies produced by these coordination actions and also the user, market and gap analysis studies in order to prepare its exploitation and dissemination plans in the search sector. In addition, links with the CHORUS / CHORUS+ project clusters will be set up when necessary in order to collaborate and re-use existing technology. |
| **VidiVideo** — Improving the accessibility of video. http://www.vidi-video.it/ The VidiVideo project focused on the development of a semantic search engine for video. The main technology behind it was the development of a 1000 element thesaurus for automatically detecting semantic concepts in the audio-visual content. | LinkedTV will exploit the technological advances made within VidiVideo to semantic concept detection in video. CERTH was a partner of the VidiVideo project, where it lead the "Video Processing" Work Package. |
| **NoTube** — Networks and Ontologies for the Transformation and Unification of Broadcasting and the Internet | While NoTube focuses on the personalisation of TV broadcasts at an atomic level (the program itself or temporal segments thereof), we will focus at a much lower granularity (individual spatio-temporal segments). Likewise, the association to related content is at an atomic level in NoTube, as opposed to LinkedTV's aim to display related media segments dynamically and in an intuitive fashion to the viewer. |

## 1.2.5. Summary of progress

It is possible to summarize the contents of this state of the art analysis, in particular comparing the expected developments in the next 4 years with the contribution that LinkedTV would bring to each area of technology.

**Table 4: Contribution of LinkedTV to relevant technology areas**

| Area of technology | Current status | Expected status in the next 4 years | LinkedTV contribution |
|---|---|---|---|
| **Web-based video** | Embedded in pages without integration of internal content to the Web | Deeper annotation and better search, increased monetization | More formal, granular (semantic) annotation leading to object-level integration of video with related Web content |
| **IPTV** | Delivery of TV over IP with back channel, STB based applications independent of current broadcast | Growth of in-TV "widgets" offering Web content parallel to and separate from the broadcast | Enabling a fine grained selection of concepts WITHIN broadcasts and browsing of related Web content parallel to the broadcast |
| **IP-based video streaming** | TV content is often replicated for the web. Users distribute video material (often of low quality) | More sophisticated publishing tools; more skilful "ordinary" producers; growth in available material | Enabling a fine grained selection of concepts WITHIN video and browsing of related Web content parallel to the video stream |
| **Video analysis and annotation** | Analysis and classification in narrow domains, with high computational cost, simple-event detection | Visual analysis and classification in progressively broader domains, and more advanced event detection | Computationally efficient analysis and classification, including instance-level object annotation and dynamic event detection, in broad domains. |
| **Integration with local data sources (home server, corporate Intranet)** | No implicit integration, only menu-based access to local data separate from the TV broadcast | Partial integration with local data by linking metadata description of atomic media items to local media | Seamless integration by rich, granular media annotation of media fragments (e.g. objects in video frames) with other media |
| **Links to external data sources** | Only a limited, dumb integration (e.g. Blinkx) with little disambiguation | Improving integration by using linguistic analysis and concept clustering to improve disambiguation | Deep integration through formal (semantic) annotation of media and unambiguous linkage to a global network of concepts (the Semantic Web) |
| **Connection to social networks (feedback, blogs)** | Only weak linking between video as a whole embedded in Web pages and data from social networks | More Web based mash-up of video material with social network data, as well as more integration of social networks into | Finer, granular linking between social network data and individual objects in the video stream |

| | | the (IP)TV platforms | |
|---|---|---|---|
| **Personalisation** | Possible on a basic level such as video genres | More refined filtering based on analysis of available data (titles, descriptions...) | Concept-centred filtering based on the object-level annotation of TV content and inference of conceptual relevance through the use of formal semantics |
| **User interfaces** | Generally "outside" of the video itself, and offering controls related to the video as a whole (e.g. EPG based program descriptions) | More intuitive interfaces for IPTV, enabling more interaction possibilities (e.g. widget menus) | Interface for the intuitive selection of individual objects in the video stream as well as the non-disruptive browsing of related media on screen |
| **Multimedia presentation** | Limited ability to automatically generate multimedia presentations without effortful manual preparation of input data | Step-wise improvements in the presentation generation through increased availability of metadata and better understanding of the rules to create presentations on that basis | A Linked Media infrastructure in the Web providing global knowledge on concepts and related media, to enable Web-based automatic multimedia presentation generation |
| **Networked Search and Retrieval** | First attempts to connect web mining with multimedia to create annotations | Use of low level video analysis to create connections between video segments and web mining results | Advanced search and retrieval based on concepts in video streams, personalization taking the social network and external resources under consideration |
| **Innovative Business Models** | Mainly different technical approaches and no focus on advertising | Commercialization by way of connection to online shops | Commercialization by way of innovative advertising models that connect to advertising in the video content and online. |

## *1.3. S/T methodology and associated work plan*

LinkedTV's overall objective is broken down into work packages (WPs) that follow the logical development of the project. The overall strategy is to perform the basic research necessary to implement a platform for future Television on the Web parallel to the architectural planning. The scenarios will also be prepared in terms of storyboards, content acquisition and development, and used subsequently to demonstrate the developed platform.

### 1.3.1 Architecture

The result of the project will be a platform which comprises all steps for the generation of meaningful linked video presentations. This covers the selection of videos, automatic annotation with metadata based on video speech and graphical objects, linking to external web sources and other videos and generation of presentations on clients with Browser or STB (Set-top box)[7].

The platform can be deployed for different soft- and hardware environments. The usage of standard interfaces and formats enables the connection of arbitrary Content Management or Media Asset Management Systems to receive video sources and to foster any other IPTV-distribution- and service platform for video transfer and user interaction management. The platform will be realized on the basis of an APS (Application Server), ESB (Enterprise Service Bus) and WebServices to connect all internal and external components, which allows to deploy flexibly other commercial or open APS for the later exploitation.

The LinkedTV platform is illustrated below:
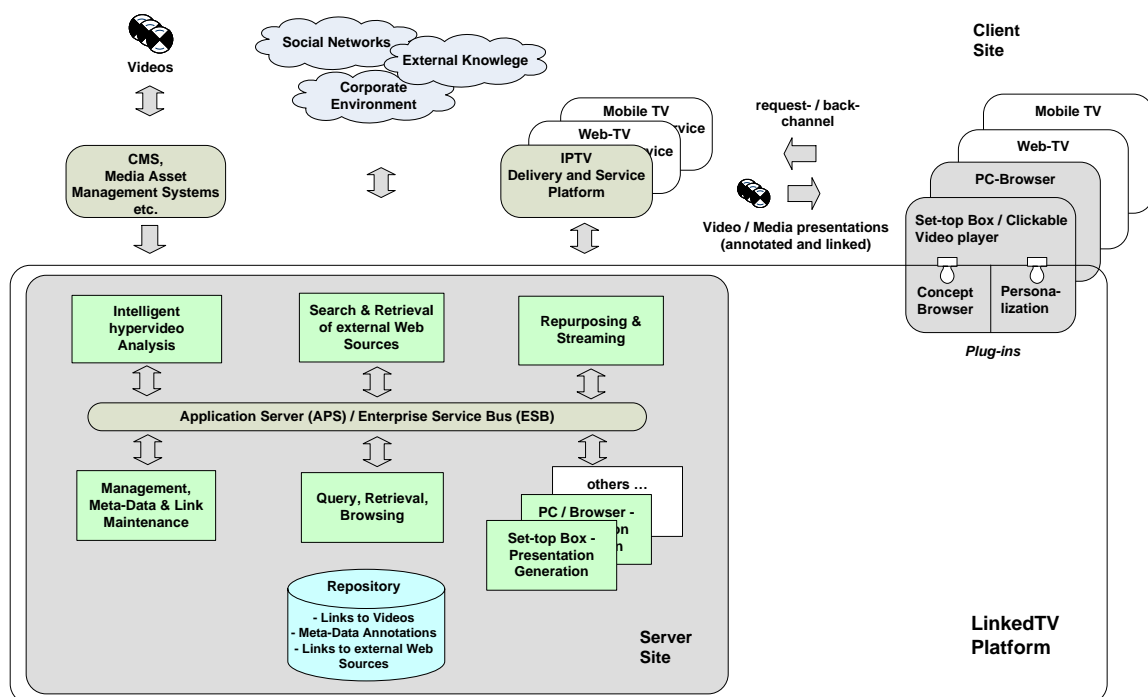


**Figure 13: The LinkedTV Platform**

---

[7] We will use a specific STB until further standardized interfaces for flexible exchange of STBs are available. There will be the options to use an open source media platform such as MythTV or STBs provided for Google-TV when available (e.g. from Sony or Panasonic).

### 1.3.2 Content side (broadcasters, archive holders, content owners)

At the broadcaster end, broadcasters can re-use their video production and preparation tools. The IT infrastructure will be extended by a means to provide rich annotation of the selected broadcasts, to store this annotation, to deliver (part of) the annotation to the end device, and to provide a query endpoint to acquire further information about broadcasts on the fly. For this, a semantic repository and tools will be developed:

- a tool for enabling content providers/owners to annotate content in a semi-automatic manner, aided by media analysis tools, data mining tools and integration with online knowledge sources (CERTH, UEP);

- a repository for storing created annotations, querying annotations and providing additional services, e.g. browsing annotations or expanding them with low level features or high level concept associations (CONDAT);

- a delivery framework that will support the broadcasting of AV material with their annotations, either interleaved in the broadcast stream or parallel by an IP channel with synchronization information (Noterik).

### Linked media: a supporting Internet infrastructure

The Linked Media infrastructure is foreseen as a decentralised Web-based platform based on the distribution of Web media annotations and associated conceptual models over the Internet (based on the Linked Data model of Web-scale interlinking of concepts and metadata about concepts, extended with media annotations which tie online media with Linked Data concepts). For the project, sample metadata stores (media annotation with concepts) will be created and linked in the Web to existing Linked Data metadata stores (concept metadata and association). The creation of the Linked Media Layer will be a project-backed initiative led by WP2 and the partners EURECOM, STI and UEP.

### 1.3.3 Service side (telecommunications operators)

For efficiency and Quality of Service, we will most likely cache Linked Media and Linked Data locally for the concepts relevant to the TV programming. This will be hosted by the service provider which can be between the broadcaster and the delivery network (injecting the service added data, such as concept viewing and browsing interface, content presentation templates and pre-prepared media presentations and other associated services such as e-shops or Web 2.0 community widgets).

We will develop

- a tool for supporting Linked Media metadata creation by enabling the service provider to browse and select media found online for a given concept, to create annotations tying online media to concepts (supported by data mining and media analysis techniques), and to index and store Linked Media metadata for efficient retrieval (EURECOM, UEP)

- programs for filtering/selecting concepts and related media on the basis of an user profile or a context description, which may run either on the broadcaster side or on the user's end device (CERTH)

- a tool for enabling service providers to select and model presentation templates for different services, concept classes, contexts etc. defining how the end user may access LinkedTV services, browse active concepts, and view sets of media associated with a concept in an appropriate and meaningful way. (CWI)

### 1.3.4 IPTV network and platform

A LinkedTV delivery network and evaluation infrastructure (testbeds) will be provided by the partner Noterik, who develops IPTV delivery solutions for the industry. The IPTV Service Platform has the functionality to

- Organize TV channels,

- Manage services,

- Configure networks,

- Configure subscribers (incl. user profiles),

- Monitor services.

It contains asset management, channel management, profile management of groups and users, plus monitoring services in a statistics service.

The LinkedTV platform will be realized on the basis of an Open Source application server with an ESB. This platform will host the modules developed by the partners. It will offer an open framework to enable each partner to develop, integrate and validate its modules independently of other partners. This framework is based on a distributed network architecture providing high scalability, high robustness and high performance. The framework integrates the management of workflows, jobs, load balancing and failover. It also integrates the supervision and the configuration of the platform via Web services, JAVA Applet and SNMP.

### 1.3.5 End user devices

The core purpose of end device development will be to provide an open multi-layered, client and server customisable canvas front-end on which all LinkedTV media and content can be projected onto, and which will be called upon using a REST API [8]. In order to play out all available media, it needs to interlink to a variety of metadata descriptors. It will be able to interface with the services in LinkedTV and will need to be able to interface to third party LinkedTV services to establish relevant mashups, and obtain high quality multimedia content from different sources such as Wikis, Maps, ASR and object tracking, for example. Research will be performed on creating an interoperable canvas, allowing smooth integration with third party XML-driven services.

The hypervideo Flash player will connect a variety of services developed within the LinkedTV project and will power the hotspot click technology on the Flash player. It will introduce video hotspot technology which will give end-users and content creators the option to find on-screen objects using layered technology enabling an enriched user experience when viewing a TV program. This component will enable direct user interaction with available content. A dense web of interacted links will steer these hotspots on the canvas. The dynamic use of manually and automatically identifying objects will enable both end-users and content creators with a state-of-the-art service which extends the video player beyond play-out. A dynamic click will enable end-users to navigate to other related content in the video itself of other available media. An API will be made available which will interact with the LinkedTV platform.

The project will also support the use of Smartphones by providing a version of the hypervideo player for Android devices. This allows to use links in TV or videos also on Smartphones with broadband connection to the internet. For porting the hypermedia player to the Android operating system, the functionality of the player will be reduced and adopted according to the

---

[8] REST (Representational State Transfer) is a software-architecture model for distributed hypermedia systems

constraints of the smaller display and input options. For the activation of links has to be considered, that on the mobile divice no mouse is available and instead multi-touch is offered by an increasing number of manufacturers. With the release of Android 3.0, which will support the input and display for tablet PCs and is announced for 2011, there will also be the option to furnish larger devices with the platform.

### 1.3.6 Work Packages

We organize the work in workpackages so that the research and development of the LinkedTV platform and its subsequent evaluation and dissemination are clearly represented in the ordering of the workpackages:

W start with workpackages that provide the **fundamental research in preparing Networked Media for interlinking across the Web**: WP1 will focus on challenges of concept-level analysis of audiovisual content, while WP2 will complement this with concept-based association to relevant Web content through data mining of Web resources. The data mining work done here will also be applied to improving the annotation from the media analysis of WP1. A suitable multimedia annotation scheme and fragment reference scheme will also be developed.

This is to be supported by workpackages designing the basis for **Television linked to the Web**: the interactive access to concepts and content through an intuitive user interface and the dynamic generation of multimedia presentations to present packages of content related to the selected concept will be tackled in WP3; while the adaptation of the concept browsing and content presentation to the user and context, whose preferences will be captured in a profile using a schema defined in the project, will be tackled in WP4.

Building on this, we have the workpackage focused on the **LinkedTV platform** (WP5): there, the integration work will take place, taking software developed from the results of the research WPs 1-4, integrating these into a backend platform with Internet search and data caching capabilities, providing network delivery capabilities (Web, IPTV) and developing the front end application for PC and STB with clickable video playback and concept/content browsing capabilities. Applying this in use case **scenarios** (WP6) will be the means to validate the LinkedTV platform functionality and achieve its transfer – prototypically – into actual WebTV and IPTV services. The scenarios will involve a test environment to deliver LinkedTV services to a selected group of testers. Through both technological evaluation (e.g. robustness, scalability) in WP5 and user trials (for usability, response levels) in WP6 we will validate the work both scientifically and socio-economically, as well as in terms of ease-of-use and quality of experience.

Finally, the prototypical use of the LinkedTV services and results of its scientific and socio-economic evaluation will be widely disseminated to both the research and industrial community. Commercial exploitation will be assured and standardization of the underlying technologies pursued. Hence, we dedicate workpackages to **dissemination** (WP7) and **exploitation** (WP8). Parallel to all activities, one workpackage will handle all **management issues** (WP9) including the regular reporting of project activities, financial matters and the monitoring of the project's progress vis-à-vis its objectives. The figure below represents the structure of the LinkedTV workpackages:
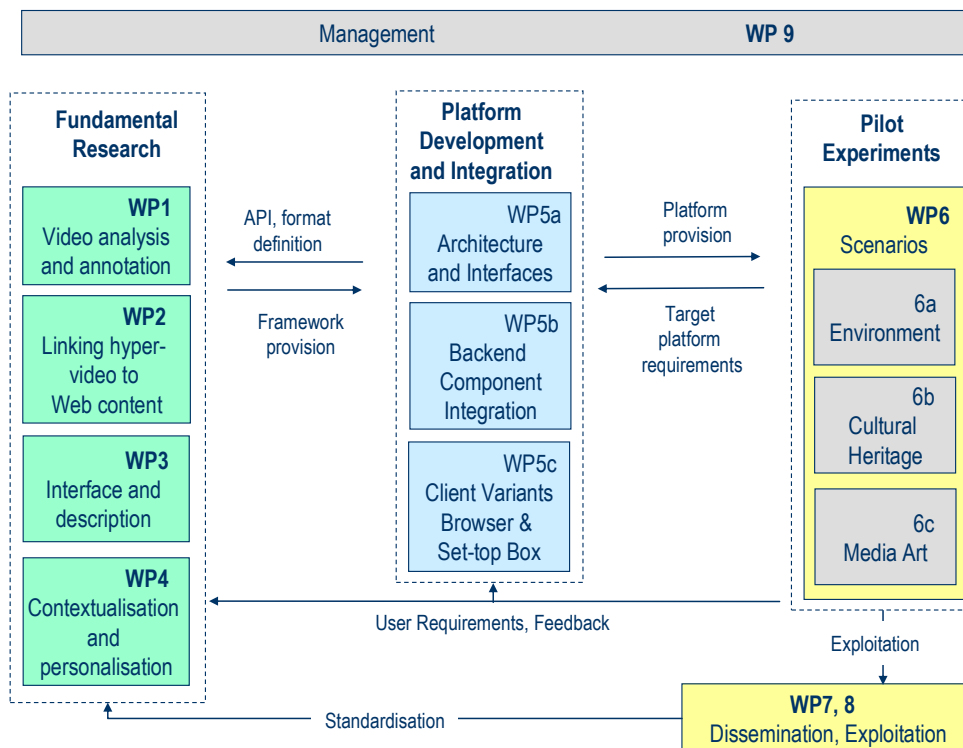
**Figure 15: WP structure**

## Detailed workplan description

*Work package 1: Intelligent hypervideo analysis*

This WP will exploit state-of-the-art tools and extend them to allow for minimal human intervention at the authoring stage and maximum user experience at the viewing stage. It will offer the technologies to *detect* content segments of interest in space and in time, *label* them with rich semantic descriptions (including event- and object-instance-labels), and finally *link* them to e.g. another part of the same or some other video, or to external sources. Overall, it addresses topics such as visual region detection, object/scene/event labeling, similarity matching, etc., as prerequisites for enabling the association between different parts of the content stream or/and also between them and other content streams or other forms of information (e.g. text, web links etc.). Specifically, this WP will provide functionalities both to the media annotators and to the end-users. The human annotator will be provided with automatically analyzed visual content, so as to minimize the time required for the manual refinement of the analysis results and ease the semi-automatic content linking process. As a result of linking, the end user will be able to select the previously annotated and linked content segments (e.g. regions), search for similar ones in the collection and receive higher level information about them (e.g. web abstracts).

*T1.1 Visual information preprocessing and representation*
This task addresses the first stages of analysis that are necessary for the envisaged automatic or semi-automatic annotation of the content. These include the automatic or semi-automatic decomposition of content to meaningful segments while considering multiple granularities (e.g. scenes, sub-shots corresponding to elementary human or non-human actions, spatiotemporal or spatial regions of interest extracted by means of automatic/interactive segmentation and tracking), the representation of these segments using SoA low- and mid-

level descriptors and extensions of them (e.g. sub-shot activity areas, region- or interest-point-trajectories over time, space-time pyramids of features, maps of facial regions), and the definition of fast similarity-based matching techniques that will exploit the developed content representations for identifying identical or similar instances of segments.

### T1.2 Visual object and scene labeling

Based on the outcome of T1.1, this task will associate content segments with a first form of semantic descriptions, i.e. an appropriate multitude of object and/or scene labels. To this end, methods that will be investigated and extended include clustering, learning (e.g. extensions of Support Vector Machines (SVM) for learning from unbalanced and subclass-structured data, techniques dealing with the time requirements of learning by exploiting GPUs or other resources), and dimensionality reduction techniques (e.g. extensions of subclass discriminant analysis). The results of this task include the representation of each content segment with a model vector (i.e. with a vector or real numbers, positioning the content in an appropriate high-dimensional semantic space of object and scene labels and taking into account the uncertainty and imperfection of each individual object detector), and the recognition of faces.

### T1.3 Complementary text and audio analysis

In this task available collateral textual information sources that can be associated directly to audiovisual content (such as subtitles, teleprompts and manual transcripts) are connected time-synchronously with the metadata. In addition, audio analysis technology is deployed to generate object segmentations based on the audio stream. Automatic speech recognition (ASR) is deployed for content for which no collateral textual sources are available. Research will focus on the recurrent, bi-directional process of using annotations from other modalities to optimize the functionality of individual annotation processes: (i) starting from ASR transcripts to suggest keywords/concepts/entities for human annotators, visual concept detectors and speaker identification, and (ii) starting from non-audio annotations or associated web-content to optimize performance of ASR (model adaptation) and speaker identification (speaker priors).

### T1.4 Event and instance-based labeling of visual information

This task will exploit the results of T1.2 and T1.3 for enriching and extending the semantic descriptions of content segments, i.e. for labeling them with event labels and for identifying specific instances of objects and events (rather than more generic object labels, as in T1.2). For the purpose of event labeling, among the methods to be investigated are elaborate combinations of dimensionality reduction, nearest neighbor classification, and learning techniques (e.g. kernel methods, not limited to SVMs), which will be applied to the model vectors of T1.2 and some of the results of T1.3 (rather than low-level visual features). For subsequently identifying specific instances of objects and events, further results of complementary text and audio analysis (T1.3) will be introduced to the labeling process. Similarly to previous tasks, where e.g. the use of GPUs or other computational resources is considered, in this task also emphasis will be put on computational efficiency (e.g. by means of dimensionality reduction), so as to successfully confront the huge amount of data that need to be processed in real application scenarios.

### T1.5 User-assisted annotation tool

This task refers to the development of a user-assisted annotation tool whose target users will be broadcasters, content owners or the LinkedTV service providers. The tool will use the combination of analysis approaches developed in tasks T1.1 to T1.4 in order to make informed suggestions to users of objects and events of interest and their labels; in this way, users should be able to complete the semi-automatic content annotation process, i.e. the linking of content, with minimal effort. As a basis, partners can provide existing tools from the FP6 projects aceMedia and LIVE for extension and improvement.

### T1.6 Intelligent hypervideo analysis evaluation

This task refers to all activities relating to the technical evaluation of the content analysis in tasks T1.1 to T1.5. These include the setup of ground truth data, the manual or semi-automatic ground truth annotation of available content according to the requirements of each task, the design of evaluation metrics and methodologies, and the performance of experiments to measure the effectiveness of the developed technologies. This task does not refer to the collection of the content itself or to the development of content annotator tools, which are addressed in different tasks or WPs. Particular focus here will be given in validating that the WP1 outcome compares favourably to existing, more conventional approaches in the field. As part of this effort, participation with WP1-developed techniques to the relevant major international benchmarking activities, such as the Semantic Indexing (SID), Instance Search (INS), and Multimedia Event Detection (MED) tasks of TRECVID, is foreseen in this task.

*Work package 2: Linking hypervideo to Web content*

This WP deals with i) the technical architecture enabling deep linking to media objects, ii) the design of lightweight metadata models, iii) the specification of a "Linked Media Layer" on the Web using this metadata in combination with Linked Data and iv) tools for mining and processing Web content in order to populate the metadata knowledge base.

Task 2.1: Media fragments addressing

The task WP2.1 aims at providing a URI-based mechanism for identifying, and retrieving the media objects detected in WP1. We will work under the assumption that an audio or a video resource has a single unified timeline. We will consider four dimensions for addressing fragments of image, audio and video resources namely: time, space, track and name. In liaison with the W3C Media Fragments Working Group, we will further specify the communication between user agents and web servers enhancing the whole web architecture for making video a first class citizen in the evolution of the world wide web. In this task, we will provide various implementations on client and server side deploying the media fragments technology.

Task 2.2: Lightweight metadata models for hypervideo

The task WP2.2 focuses on developing lightweight models for metadata interoperability. Media objects automatically detected in WP1 will be annotated using simple metadata schemas while being compatible with the numerous multimedia metadata standards. The starting point will be the COMM multimedia ontology developed within the FP6 EU project K-Space. We will adapt and simplify this model in order to create a Lite version according to the needs of the LinkedTV use cases. We will implement an API and the necessary interfaces to be integrated in the LinkedTV infrastructure. In liaison with the W3C Media Annotations Working Group, we will ensure that the metadata schema is largely interoperable with the numerous metadata standards, providing mappings between the description properties.

**Task 2.3: Web mining for support of media annotation**

This task exploits web content relevant to the broadcast to assist with the identification and annotation of semantic concepts and objects as carried out in WP1. WP1 performs low-level analysis of a multimedia object while exploiting complementary audio and text directly associated with the audiovisual content. Using information mined from the web can help further leverage the associated content. For example, many entities appearing in the complementary resources such as transcripts or subtitles are named entities (names of people, places, organizations, etc.) that are without external knowledge or context "meaningless" to the machine. However, once additional information obtained from the web is provided, these

named entities can be resolved and the information conveyed by them unfolded. Additionally, they become unambiguous and hence can provide a hint for disambiguation of other entities.

The primary methods employed in WP 2 will include wrapper-based extraction from structured web content and Named Entity Recognition and entity classification over free-text web content (scripts, subtitles, online reports to sports events etc.), or a mix of these (e.g. from programme guides).

Novel methods of term expansion and disambiguation using WordNet, DBPedia, the structured version of Wikipedia, and domain-specific databases will be developed, in order to establish the right context. We will benchmark and compare numerous Information Extraction tools such as OpenCalais, SPROUT, GATE, KIM and evaluate novel disambiguation algorithms.

**Task 2.4: Retrieving additional content from the web**

The goal of this task is to retrieve and pre-process external web resources to enrich the broadcasts with additional information about important concepts that could be presented to the viewer. We will investigate the use of advanced techniques for web mining that have not yet been considered in connection with multimedia concepts. First, large collections of web resources, forming an 'information cloud' around individual broadcasts, will be partly manually pre-selected and partly retrieved via intelligent meta-search. Subsequently they will be examined at the level of individual website and document, respectively, using web spiders and document categorisers. Finally, concrete documents selected via spidering and classification will be submitted to fine-grained information extraction (IE) tools.

Since there is a limited number of different broadcast genres and assuming that the viewers are likely to have different preferences and requirements for each of the genre, there is a need for creating genre-specific information gathering templates. These templates would provide the necessary granularity and adaptability to the user's requests and interests.

As core IE techniques for this purpose we envisage rapid prototyping of rich extraction models combined with partial statistical model training (offline) and local wrapper induction (on the fly). The structure of the extraction models could borrow from the structure of Linked Data resources (such as DBPedia) in which the source concepts are anchored. The hybrid approach (rich hand-crafted extraction model + trained statistical model + local wrapper) is particularly suitable for collections of semi-structured, field-specific web resources, whose internal structure can only be roughly estimated. Together with textual resources, structured resources with similar scope will be exploited. For example, extraction from free-text Wikipedia articles will be complemented with and/or triggered by processing of DBpedia. Furthermore, available social network metadata will be used. Text processing techniques such as co-occurrence analysis and named entity recognition will be applied to loose metadata structures such as folksonomy tags.

*Work package 3: LinkedTV interface  and presentation engine*

The work will be carried out in close cooperation with WP1 and WP2 on the underlying knowledge being made available, with the personalization and contextualization work in WP4, and with the front-end being developed in WP5 and the implementation of the scenarios in WP6.

The workpackage will construct hypervideo interaction interfaces for supporting information browsing, organization and presentation, and higher level tasks such as information gathering. It is based on the tasks implicit in the 3 scenarios and will generalize results to identify design patterns and guidelines.

The interface design will strive to hide the complexity of the LinkedTV system operating in the background. For example, the user should be unaware of the connection between the video pixels and the concept describing them, the relation between this and a thesaurus, and how these are enriched with other web-based information. At the same time, the user should be allowed access to the underlying information resources to enable them to check information sources when needed. A metric for success for the LinkedTV interfaces is that the user should not even notice "the interface", but that they should be able to carry out their tasks with a minimum of effort.

Task 3.1 User requirements analysis

A list of information tasks related to hypervideo will be developed in the initial stages of the project. These will be selected from tasks that occur in more than one of the scenarios, such as, e.g., highly visual information browsing (e.g. for children's edutainment interfaces and cultural heritage), orinformation gathering for professional information users (e.g. environmental issues and cultural heritage). From the list of identified tasks, a small number (3 or 4) will be selected for design, implementation and evaluation within the project. The task will also collect requirements for the interfaces to be developed. Examples of such requirements are:

- Users engaged in audiovisual content should not be overly distracted.
- The role of the social context and the personal interests of the user should be taken into account.

The tasks will relate directly to the three use cases and take into account the information sources being made available in WP1 & WP2. A number of tasks will be selected (in collaboration with content partners) for which innovative interfaces will be created in Task 3.2.

Task 3.2 User-centred design and specification of hypervideo information interfaces

Per selected task, a user-centred approach will be taken to creating at least two, preferably more, different designs of each interface. These will be discussed with potential users, and content providers in the project. Out of this process final interfaces will be selected for development and inclusion in the LinkedTV platform. A specification for interface description will be decided upon whose abstract description model is flexible enough to allow for alternative user interfaces and their adaptability according to user and context. Examples of potential interfaces to be developed are:

- support for browsing/novel discovery of long tail content;
- a visual navigation interface allowing the browsing of related images/video thumbnails while not having to type;
- an information gathering interface while allowing a user to continue browsing;
- an interface for information professionals to allow them to compare pieces of information for making decisions.

**Task 3.3 Develop the interface and presentation engine**

This task will develop the engine which can generate individual user interfaces on the fly according to the current user, context and material being consumed, and deliver those interfaces to the front-end device for adaptation and display.

**Task 3.4 Evaluate functionalities and interfaces of the connected media and background information**

Evaluations of the interfaces included in the LinkedTV platform will be carried out. This will be done with potential users of the LinkedTV platform, selected through the content/broadcast

partners in the consortium. This will complement the evaluation of the end-to-end system in WP5 and the implementation of the scenarios in WP6.

The results of the evaluation will be generalized as design guidelines that can be used in the creation of interfaces in other scenarios and identifying hypervideo interaction design patterns where possible.

*Work package 4: Contextualisation and personalisation*

**Objectives**

This WP deals with user tracking and profiling, with respect to activity and behaviour processing, in order to provide ConnectME end-users with effective content filtering based on their preferences and interests. It will exploit state-of-the-art tools and approaches, and extend them to allow for design, development and testing of intelligent methodologies and technologies for user tracking and the subsequent personalization of multimedia content, in terms of both object selection and content presentation. Intelligent social and content-based personalization methodologies will be implemented based on domain-independent personalization strategies that will elaborate recommendations based on the semantic relationships and context between the user's preferences and the available content items.

*T4.1 User profiling*

This task deals with the processing and extraction methods involved in the modelling and extraction of ConnectME profiling information. It will enable mining of profile elements through user interaction and exploitation of such personal information to provide more personalised access to the content.

More specifically, user preferences will be captured unobtrusively based on the user's interactions with the video, in addition to any user-defined explicit interests (including integration of Web-based information about interests derivable from the user's profiles on social networks). The concepts extracted from the consumed content (WP1 and WP2) will be accumulated/updated to the user profile.

Utilizing recent advances in the information foraging theory and Multiple Criteria Decision Analysis, the history of user's interaction with content can be used to understand, rather than only learn, his or her preferences. The resulting model of information scent or utility curves can then be stored in the visitor's profile and used for contextualization.

User interests will be expressed in formal semantics (based on the information provided by WP2), along with their degree of participation to the profile, in order to form a meaningful user profile. A temporal decay factor will be used to update concept weights in order to discern between short and long term interests.

*T4.2 User Behavior and Contextualization*

This task addresses the issues of modeling a strategy to anticipate user actions and provide focused interaction, by monitoring the user's physical interaction with the video. A set of behavioral features will be extracted from sensors such as cameras, 3D depth cameras (time-of-flight), or microphones on user behavior including the focus of attention (gaze detection, quantity of motion, contraction index, symmetry index, voice signals, etc…), as well as collective reactions. Collective reactions may include collective laughter, questions that children ask their parents following a popular science program, or lively arguments in front of controversial messages. Those innovative techniques around user voice, eye and body tracking will be used to enhance the personal profiles and preferences defined and extracted within T4.1. Current state-of-the-art personalized information filtering, ranking and

adaptation techniques will be extended, using the tools of T4.1 and the profile and knowledge information extracted in this task.

Besides detecting predefined behavior patterns, such as clear expressed emotions, and attention focus on particular object categories, the approach to behavior tracking will rely on the computation of a rarity index of the tracked features on either a short period of time (some seconds) or on more long-term periods is able to highlight novel/surprising behavioral patterns on one or more of those features. An unusual reaction is likely to exhibit a novel emotional reaction that the user experienced during his/her video browsing or retrieval activity. Information about the similarity of reactions of the current user with that other users for the same video may be a cue that there are similarities also between the kind of media they like, thus the profile of the other users can help in refining the research of the current user.

The extraction of local features in the regions of interest (detected either automatically or by gaze tracking) provide the system with the possibility to highlight the multimedia files which exhibit a similar local descriptors configuration or to highlight a particular object in the video or present a particular piece of information to the user. In relation to WP3, this can be used as trigger to prompt manual annotations of the media object.

At the group level, when several people have been watching the same video, averaging can provide an emotional evolution profile for the content. In that case, it is also possible to detect outlier behavior, where people react in a different manner than the majority of the users.In relation to WP3, this can be used to prompt people comments on the current media in order for him to express his potentially different opinion and to provide ratings.

Besides providing information for the personal profile, this approach can also be used to enrich the media description in terms of behavioural responses elicited by the media or media fragment.

### T4.3 Personalized content filter

This task will focus on providing algorithms for filtering and suggesting concepts and content based on the user's interest profile, and previous activity and behaviour. This will be achieved by semantically matching the user profile against the available linked content (semantic representation) with respect to the knowledge in context, while taking under consideration the weight of the user preferences in order to rank proposed concepts and content. Fast and efficient algorithms will be provided, aiming to integrate the filterer on any end-user device. In case of scarcity of available linked content, a user model will be provided in order to formulate a search query for "free" web search.

### T4.4 Evaluation of personalization technologies

This task refers to all activities relating to technical evaluation of contextual and personalized content processing and information fusion in tasks T4.1, T4.2 and T4.3. These include setup of ground truth data and test-bed user communities, according to the requirements of each personalization task. It also includes design and implementation of meaningful evaluation metrics and methodologies, as well as performance of experiments to measure the effectiveness of the developed technologies.

### Work package 5: LinkedTV platform

WP 5 is the main work package which ensures that the result of LinkedTV will be a highly integrated, open, and manageable end-to-end platform. This includes the following objectives:

- Defining an architecture that integrates the components for video/audio and Web content analysis, metadata generation and storage, creation of clickable video and user interfaces, and their delivery and presentation to the user into an end-to-end platform
- Defining APIs and protocols to connect all components and data flows.
- Implementing interfaces, data transfer and input/output functions not provided by WPs 1-4
- Integrating and testing of all components designed and implemented in WPs 1-4
- Coordinating the other technical WPs 1-4: WP5 will ensure that the algorithms and technological solutions for the back-end processing work together smoothly and error-free.
- Developing the front-end system for PC and STB which provides the clickable video player and LinkedTV user interface, ensuring interoperability with the back-end data and processes
- Coordinating and self-assessing the use case implementations in WP 6 by fixed synchronisation points with the use cases and an iterative self-assessment task

To achieve this, WP5 has three primary sub-responsibilities:

5a:  Development and use of an integrating platform for the end-to-end functionality

In addition to the components developed in WPs 1-4, we will ensure the preparation and inclusion of back-end functionality for the storage and retrieval of the generated audiovisual material annotations as well as a search and retrieval component for accessing Web (meta)data for concepts and related content. Therefore, the interfaces to be used by backend components as well as front-end clients for accessing the application platform have to be specified here.

CONDAT will provide the integration platform for development, and delivery modules will be provided by Noterik for video and data delivery over IP networks, both to PC-based Web browsers as well as a STB-based IPTV application.

The architecture will be built as far as possible on existing or emerging standards and reliable open source frameworks and products, especially results from other EC co-funded projects For the validation of the correct cooperation of all components and performance assessment a representative set of benchmark scenarios will be developed with respect to the use cases of WP6. The resulting platform will build an open service framework for the implementation of the use cases and further value added services of 3$^{rd}$ party providers.

5b: Backend component integration and testing of all back-end components to provide the full back-end functionality

The components realized in WP 1-5 will be integrated in three cycles. The interfaces for the integration of all components will be defined flexible enough, so that a repeated integration during the 2 feedback-development cycles will be possible. The services for browsing, search and retrieval for the users on both types of clients will be implemented and provided for the use in the trials.

After the successful integration of all components on the back-end and client side an evaluation will be performed. This covers firstly verification, proving that the implementation is according to the functional specification. This guarantees a correct cooperation of all components and performance assessment by using a representative set of benchmark scenarios. Secondly a validation is performed, analysing whether the provided results are accepted by the end users. The final validation will be done during the trials in WP 6.

5c: Development and installation of the client variants

The front-end playout will be enabled by the development of a Flash-based  MediaCanvas API and its support for clickable video technology which enables the layering over this video of further interactive regions allowing for the presentation of the LinkedTV user interfaces. This specific component will enable direct user interaction with available content. A dense web of interacted links will steer these hotspots on the canvas.

The front-end needs to be interactive and fast and therefore it will be providing LinkedTV with a caching technique which will enhance not only a fast user experience, but aims to provide a scalable adaptive caching mechanism. The adaptation is needed for the terminals the playout is aimed at and will scale accordingly. In order to ensure this, research will be performed on a pre-caching mechanism, whereby the caching mechanism will establish first playback option before the actual caching starts. Research will also be performed onto real-time caching on fresh content which may be added to a content repository without the caching mechanism serving cache-only content.

## Work package 6: Scenarios

This workpackage is made up of three use cases which will be realized in parallel. Although the use cases and scenarios deal with different topics and content areas, the work to be undertaken in principal will be similar in nature. Hence, the individual tasks are the same, differing only in their content focus. The work package will take care to prepare effectively for the realization of the scenarios by identifying and collecting the media and the metadata to be used, as well as to carry out the actual realization on top of the LinkedTV platform provided in WP5. Finally, evaluations of the scenarios will be organized through internal testing, user trials and public feedback gathering.

6a: Environmental

Environmental issues are growing in importance day by day. They have many facets, ranging from their economic implications (creating jobs, becoming a growing economic sector etc) to their importance for mankind as a whole. The LinkedTV consortium selected a scenario dealing with environmental issues as it can be certain that these topics will remain on the world's agenda in the years to come. Hence there will certainly be a need for both: making access to related content easier and more user-friendly, and dealing with better presentation (e.g. through cross-linking etc). However, the approach selected is such that, once it is working in the environmental field, the system and underlying technology can easily be ported to other domains.

6b: Cultural Heritage

Cultural heritage is by nature strongly interlinked, e.g. thematically, historically, and geographically, both within itself and to current events. In order to fully understand the events of today, knowledge about the historical background is needed. For example, the idea of the European Union is related to the history of Europe, e.g. the numerous wars and constant rivalry between the European countries and, on the other hand, as a response to global competition and policies. This usage scenario aims at providing the viewer the cultural heritage context of currently viewed events, actors (persons, organizations, etc.), places and other relevant aspects.

6c: Media art

Broadcasters, content producers, but also artists, need to find new ways to include Web-based content into the experience to make it more entertaining, educational, or appealing to the senses or the intellect. The media art scenario focuses on the experimentation in how an advanced multimedia service, based on LinkedTV tools, may allow to empower artists and content creators with stronger means for content reuse, be it for inspirational purposes, or for creating content mashups, as well as with better ways for considering public feedback.

*Work package 7: Dissemination*

One of the aims of the LinkedTV project is to disseminate information about the project, its objectives, the approaches chosen and its results in a professional and high quality manner, using a variety of means and channels.

The consortium will make sure that LinkedTV activities are disseminated to various communities and interest groups. These include potential end customers of future LinkedTV products and services,potential future industrial collaborators, the European Commission and its research community (via project officers, other ICT projects in related technological areas ,with which clustering activities will be established), other researchers and technology developers in academic institutes and industrial research laboratories.

A LinkedTV project web-site will be set up. It will be updated regularly with information for the general public, representing potential end users and broadcasters, content owners, content providers and network operators - who represent the technology adopters - and provide a restricted area for project partners. A publication plan will be developed and an overview of exhibits and relevant conferences will also be provided in order to make sure project partners know well ahead of relevant events taking place. Special focus will be made on showcasing LinkedTV technologies in an exciting and interactive way, both by allowing for interactive video on the website as well as demonstrators at chosen research and industry events.

The objectives of LinkedTV, the approaches chosen, and the respective results will be made available on the website as well. We will create a two-way communication channel with stakeholders, standardization bodies, academic communities and industry for disseminating results. Apart from this taking place via the project website, this will also be done through Web 2.0 channels, e.g. social networks, blogs, video sharing sites, tweets and the like.

Activities towards standardisation of the project results will also be explored and coordinated in the Dissemination workpackage, so that LinkedTV will have the best possible impact both in the scientific and commercial communities. In terms of current standardization efforts, LinkedTV will actively participate in and contribute to various standardisation bodies such as W3C, IPTC and EBU, with the expectation of having significant impact on their development. Consortium partners are already participants and co-chair specific technical working groups within these bodies.

In sum, following dissemination activities will be undertaken:

- Project Website
  - Information about project (public), also using features such as RSS feeds etc
  - Shared work space (private)
- Production of dissemination materials
  - Brochures
  - Posters
  - Videos / films

- Animations (e.g. usage scenarios)
- Web based dissemination
    - Video uploads/viral video
    - Blogging, tweeting
    - Social network groups
- Market watch (products and services), including an overview of relevant conferences, workshops and trade fairs etc for dissemination
- Participation in clustering activities initiated by EC
- Participation at conferences, seminars, trade fairs etc
- "Spreading the word" through existing networks
- Publications in scientific / trade journals etc
- Pre-standardization and standardization activities

## *Work package 8: Markets, Business Models and Exploitation Strategies*

This work package will focus on the markets where LinkedTV can have an impact, by providing a market analysis, developing effective business models for LinkedTV services and exploitation strategies for LinkedTV technologies in order to lay the path for a successful technology and knowledge transfer from the project to industrial players interested in the commercialisation of the results. One goal of this WP is to provide an overview of current developments and trends on the IPTV and Web TV market based on a broad market analysis. A second goal of this WP is assessment of user requirements based on qualitative and quantities empirical research. Finally based on the results from the first two activities, the third goal of this WP is the development of innovative business models from the perspective of the most important players involved in the value chain: content provider, IPTV provider and advertising companies. This activity covers furthermore the effort made towards the definition of the partner's exploitation plans including the definition of a roadmap to exploit, use and disseminate the relevant final results.

### T8.1 Product & Market Surveys

LinkedTV will develop an innovative platform that has the potential to change considerably how IPTV and WebTV are consumed. Technically there are many options which TV and Internet information to connect as well as how to create the user experience. However, due to the high innovativeness of the project there is little knowledge about the specific user needs and preferences regarding the way how and which media should be connected. For example the knowledge which content connection are preferred by users as well as in which form might help to make technical choices. In addition, there is little systematic knowledge about the requirements and preferences of content providers. For example, are there any preferences of content providers which part of the content to connect? How does LinkedTV impact advertising provided with the content? What might be the allowed and preferred modes of connection and interaction from the content provider perspective? How can the content provider profit more by allowing content to be used by LinkedTV? Finally, an open question if and in which form advertising might get included in LinkedTV business models?

In order to answer the above questions a broad market analysis will be conducted among potential users and players in the value chain - in particular content and IPTV providers as well as advertisers.

Activities:

- **Overall market analysis** about market potential and trends based on secondarily market data available. The overall market analysis should provide a detailed overview of:
    - the state-of-the-art market developments and future opportunities;
    - of existing competitors
    - and strengthen the argumentation for the USP of LinkedTV.
- **Market analysis of user needs** based on quantitative and qualitative market analysis instruments. The goal of this analysis is to find out which content connection are preferred by users and what are the preferred modes of interaction, Is the user willing to pay for connecting content and better experience? The results of this activity should be input to use case development and also architecture.
- **Market analysis of content and IPTV provider** needs, based on expert interviews and other quantitative and qualitative market research instruments. The goal of this activity is to find out what a plausible way of connecting content would be for content providers. What kind of new advertising models might be possible and which are preferred.

The right mix of preferred functionality will be at the intersection of the preferences of users, content and IPTV providers. The result will be input to development of use cases and functional requirements. It will also assure a user centered development.

## T8.2 Business Models

The innovative technology LinkedTV has the potential to change existing content and advertising value chains for IPTV and WebTV. On the one hand new relationships among the players in the value chain might arise as well as new power positions on the market. On the other hand new business models might get possible for all players involved. For example: the content of the content providers is the starting point to creating clicks and even initiating transactions on the web. This means that on the one hand LinkedTV allows content providers to tap into existing advertising chains online. on the other hand new relationships and new forms of advertising might become possible.

The goal of this workpackage is the analysis of the new positions and options of all involved players and development of innovative business models and value chains that provide win-win and revenue possibilities for new players.

Activities:

- Analysis of value chains
- Conception of business models based on the MCM business model framework.

## T8.3 Exploitation Planning & Exploitation Activities

The goal of this task is a clearly envisioned market entry strategy for LinkedTV technologies, consortium partners and other players in the value chain as well as a clear vision how the project results will be turned into a commercial success. In order to assure effective preparation of exploitation very early in the project an exploitation plan will be developed containing the goals of the partners and serving as a guideline for activities during the project. LinkedTV offers business opportunities for all involved partners in the project and all involved players in the LinkedTV value chain. Within this activity besides development of individual and common exploitation strategies the plan will also include concrete scheduling of presentation of project results to companies involved in the project and other companies that can take commercial advantage of the project results. This final version of the exploitation strategies will be based on the results from the market analysis and developed concepts for business models. As far as possible exploitation plans will include quantification of opportunities for partners and other players in the value chain.

**Activities:**

- Development of exploitation overall plan for exploitation activities during the project
- Development of individual and common exploitation plans for each partner and involved players and as far as possible quantification of opportunities. The exploitation strategies might vary in scope and goal. We expect exploitation strategies by way of inclusion of LinkedTV developments in existing products of involved partners, by reusing methodologies and software components, by creating new content formats and others.

*Work package 9: Management*

The Management work package is concerned with ensuring that:
- the project remains on course,
- it is effectively and correctly managed financially,
- the work plan defined in this document is adhered to or adjusted if necessary,
- project progress and status are efficiently and effectively monitored,
- the required reporting is prepared and delivered in an accurate and timely manner,
- all quality assurance and self-assessment aspects of the project are fully and correctly addressed,
- the infrastructure supporting the Web-based facilities to be used for dissemination and central intra-project communication and cooperation are managed properly.

Moreover, tight and careful contingency planning and controlling activities are vital for the success of the project. The relevance of the risks are continuously checked and updated in the course of the project reflected in periodic reports. Work package leaders and board members of the management bodies will periodically report, assess and update the progress of the project as defined in the section about project management. This work package particularly includes public documentation and reporting as required by the European Commission.

**Table 4 : Work package list**

| Work pack-age No[9] | Workpackage title | Type of activity [10] | Lead partici-pant no[11] | Lead participant short name | Person month-s[12] | Start[13] | End [14] |
|---|---|---|---|---|---|---|---|
| 1 | Intelligent hypervideo analysis | RTD | 2 | CERTH | 150 | 1 | 42 |
| 2 | Linking hypervideo to web content | RTD | 5 | EURECOM | 140 | 1 | 42 |

---

[9] Work package number: WP 1 – WP n.

[10] Please indicate <u>one</u> activity per work package:

RTD = Research and technological development (including any activities to prepare for the dissemination and/or exploitation of project results, and coordination activities); DEM = Demonstration;

MGT = Management of the consortium; OTHER = Other specific activities, if applicable in this call.

[11] Number of the participant leading the work in this work package.

[12] The total number of person-months allocated to each work package.

[13] Measured in months from the project start date (month 1).

[14] Measured in months from the project start date (month 1).

| 3 | LinkedTV interface and presentation | RTD | 4 | CWI | 104 | 1 | 42 |
|---|---|---|---|---|---|---|---|
| 4 | Personalisation and contextualisation | RTD | 1 | FRAUN-HOFER | 131 | 1 | 42 |
| 5 | LinkedTV platform | RTD | 8 | CONDAT | 87 | 1 | 42 |
| 6 | Scenarios | RTD | 7 | STI | 102 | 1 | 42 |
| 7 | Dissemination | OTHER | 7 | STI | 72 | 1 | 42 |
| 8 | Exploitation | RTD | 6 | USG | 77 | 1 | 42 |
| 9 | Management | MGT | 1 | FRAUN-HOFER | 33 | 1 | 42 |
|  | TOTAL |  |  |  | 896 | 1 | 42 |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |

**Table 5: List of Deliverables[15]**

| Del. No[16] | Deliverable name | WP no. | Nature [17] | Dissemi nation level[18] | Delivery date[19] (proj. month) |
|---|---|---|---|---|---|
| D7.1 | Project website | 7 | R | PU | 2 |
| D2.1 | Specification of the Connected Media Layer | 2 | R | PU | 6 |
| D3.1 | Requirements document | 3 | R | PU | 6 |
| D3.2 | Interface and presentation specification | 3 | R | PU | 6 |
| D4.1 | Specification of user profiling and contextualisation | 4 | R | PU | 6 |
| D5.1 | LinkedTV platform and architecture | 5 | R | PU | 6 |
| D7.2 | Dissemination and standardisation plan | 7 | R | PU | 6 |
| D8.1 | Exploitation plan for the project | 8 | R | CO | 6 |
| D1.1 | State of the art and requirements analysis for hypervideo | 1 | R | PU | 12 |
| D2.2 | Specification of media fragments and multimedia metadata | 2 | R | PU | 12 |
| D2.3 | Specification of Web mining for hypervideo concept identification | 2 | R | PU | 12 |
| D3.3 | LinkedTV user interfaces | 3 | R | PU | 12 |
| D3.4 | Interface and presentation engine | 3 | R | PU | 12 |
| D4.2 | User profile schema and profile capturing | 4 | R | PU | 12 |
| D4.3 | Content and concept filter v1 | 4 | R | PU | 12 |
| D5.2 | LinkedTV front-end: video player and MediaCanvas API | 5 | P | PU | 12 |
| D6.1 | Scenario descriptions | 6 | R | PU | 12 |
| D8.2 | First market analysis | 8 | R | PU | 12 |

---

[15]  where appropriate and possible (such as for the research tasks), deliverables will be research publications (e.g. submissions to conferences or journals).

[16] Deliverable numbers in order of delivery dates. Please use the numbering convention <WP number>.<number of deliverable within that WP>. For example, deliverable 4.2 would be the second deliverable from work package 4.

[17] Please indicate the nature of the deliverable using one of the following codes:
     **R** = Report, **P** = Prototype, **D** = Demonstrator, **O** = Other

[18] Please indicate the dissemination level using one of the following codes:
     **PU** = Public
     **PP** = Restricted to other programme participants (including the Commission Services).
     **RE** = Restricted to a group specified by the consortium (including the Commission Services).
     **CO** = Confidential, only for members of the consortium (including the Commission Services).

[19] Measured in months from the project start date (month 1).

| D1.2 | Visual, text and audio information analysis for hypervideo, first release | 1 | R | PU | 18 |
|------|------|---|---|----|----|
| D5.3 | First LinkedTV end-to-end platform | 5 | P | PU | 18 |
| D7.3 | Dissemination and standardisation report v1 | 7 | R | PU | 18 |
| D8.3 | LinkedTV business models, v1 | 8 | R | PP | 18 |
| D1.3 | LinkedTV annotation tool, first release | 1 | R | PU | 24 |
| D2.4 | Annotation and retrieval of media by concept in the Web | 2 | R | PU | 24 |
| D2.5 | Final LinkedTV metadata schema and mappings | 2 | R | PU | 24 |
| D3.5 | Requirements document v2 | 3 | R | PU | 24 |
| D3.6 | Interface and presentation engine v2 | 3 | R | PU | 24 |
| D4.4 | User profile and contextual adaptation | 4 | R | PU | 24 |
| D4.5 | Content and concept filter v2 | 4 | R | PU | 24 |
| D5.4 | Final LinkedTV integrating platform | 5 | P | CO* | 24 |
| D6.2 | Scenario demonstrators | 6 | D | PP | 24 |
| D7.4 | Project demonstrator v1 | 7 | D | PU | 24 |
| D8.4 | Evaluation of exploitation impact and recommendations | 8 | R | PU | 24 |
| D8.5 | LinkedTV end-to-end platform value chain | 8 | R | PU | 24 |
| D5.5 | LinkedTV front-end: video player and MediaCanvas API v2 | 5 | P | PU | 30 |
| D6.3 | User trial results | 6 | R | PU | 30 |
| D7.5 | Dissemination and standardisation report v2 | 7 | R | PU | 30 |
| D1.4 | Visual, text and audio information analysis for hypervideo, final release | 1 | R | PU | 36 |
| D1.5 | LinkedTV annotation tool, final release | 1 | R | PU | 36 |
| D2.6 | Advanced concept labelling by complementary Web mining | 2 | R | PU | 36 |
| D3.7 | LinkedTV user interfaces v2 | 3 | R | PU | 36 |
| D4.6 | Contextualisation solution and implementation | 4 | R | PU | 36 |
| D5.6 | Final LinkedTV end-to-end platform | 5 | P | PU | 36 |
| D6.4 | Scenario demonstrators | 6 | D | PU | 36 |
| D7.6 | Project demonstrator v2 | 7 | D | PU | 36 |
| D8.6 | Market and product survey for LinkedTV services and technology | 8 | R | PU | 36 |
| D8.7 | LinkedTV business models, v2 | 8 | R | PP | 36 |

| D1.6 | Intelligent hypervideo analysis evaluation, final results | 1 | R | PU | 42 |
|------|------------------------------------------------------------|---|---|----|----|
| D2.7 | Final Connected Media Layer and evaluation | 2 | R | PU | 42 |
| D3.8 | Design guideline document for concept-based presentations | 3 | R | PU | 42 |
| D4.7 | Evaluation and final results | 4 | R | PU | 42 |
| D5.7 | Validation of the LinkedTV architecture | 5 | R | PU | 42 |
| D6.5 | Final evaluation | 6 | R | PU | 42 |
| D7.7 | Dissemination and standardisation report v3 | 7 | R | PU | 42 |
| D8.8 | Common and individual exploitation plans for after the project | 8 | R | PP | 42 |

# Gantt Chart

**Table 6: Workplan**

| WP | WP and Task Descriptions | Start (Mo) | End (Mo) | Dura-tion (Mo) | Year 1 (M1–M12) | Year 2 (M13–M24) | Year 3 (M25–M36) | Year 4 (M37–M42) |
|---|---|---|---|---|---|---|---|---|
| 1 | **Intelligent hypervideo analysis** | 1 | 42 | 42 | ████ | ████ | ████ | ████ |
| T1.1 | Visual information preprocessing and representation | 1 | 24 | 24 | ▓▓▓▓ | ▓▓▓▓ | | |
| T1.2 | Visual object and scene labeling | 1 | 36 | 36 | ▓▓▓▓ | ▓▓▓▓ | ▓▓▓▓ | |
| T1.3 | Complementary text and audio analysis | 1 | 24 | 24 | ▓▓▓▓ | ▓▓▓▓ | | |
| T1.4 | Event and instance-based labeling of visual information | 7 | 36 | 30 | (M7–) | ▓▓▓▓ | ▓▓▓▓ | |
| T1.5 | User assisted annotation tool | 7 | 36 | 30 | (M7–) | ▓▓▓▓ | ▓▓▓▓ | |
| T1.6 | Intelligent hypervideo analysis evaluation | 13 | 42 | 30 | | (M13–) | ▓▓▓▓ | ▓▓▓▓ |
| | | | | | | | | |
| 2 | **Linking hypervideo to web content** | 1 | 42 | 42 | ████ | ████ | ████ | ████ |

| | | | | | |
|---|---|---|---|---|---|
| T2.1 | Media fragments addressing | 1 | 36 | 36 | |
| T2.2 | Lightweight metadata models for hypervideo | 1 | 36 | 36 | |
| T2.3 | Web mining for support of media annotation | 1 | 42 | 42 | |
| T2.4 | Retrieving additional content from the web | 13 | 42 | 30 | |

| WP | WP and Task Descriptions | Start (Mo) | End (Mo) | Dura-tion (Mo) | M1 | M2 | M3 | M4 | M5 | M6 | M7 | M8 | M9 | M10 | M11 | M12 | M13 | M14 | M15 | M16 | M17 | M18 | M19 | M20 | M21 | M22 | M23 | M24 | M25 | M26 | M27 | M28 | M29 | M30 | M31 | M32 | M33 | M34 | M35 | M36 | M37 | M38 | M39 | M40 | M41 | M42 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **3** | **LinkedTV interface and presentation engine** | 1 | 42 | 42 | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ |
| T3.1 | User requirement analysis | 1 | 6 | 6 | ░ | ░ | ░ | ░ | ░ | ░ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T3.2 | User-centered design and specification of hypervideo information interfaces | 1 | 36 | 24 | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | | | | | | | | | | | | | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | | | | | | |
| T3.3 | Develop the interface and presentation engine | 7 | 24 | 18 | | | | | | | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | | | | | | | | | | | | | | | | | | |
| T3.4 | Evaluate functionalities and interfaces of the connected media and background information | 19 | 42 | 12 | | | | | | | | | | | | | | | | | | | | ░ | ░ | | | | | | | | | | | | | | | | | ░ | ░ | ░ | ░ | ░ | ░ |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| **4** | **Contextualisation and Personalisation** | 1 | 42 | 42 | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ | █ |
| T4.1 | User profiling | 1 | 24 | 24 | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | | | | | | | | | | | | | | | | | | |
| T4.2 | Contextualisation | 1 | 36 | 36 | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | | | | | | |
| T4.3 | Personalized content filter | 7 | 36 | 30 | | | | | | | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | | | | | | |
| T4.4 | Evaluation of personalisation technologies | 37 | 42 | 6 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | ░ | ░ | ░ | ░ | ░ | ░ |

| WP | WP and Task Descriptions | Start (Mo) | End (Mo) | Duration (Mo) | Year 1 | | | | | | | | | | | | Year 2 | | | | | | | | | | | | Year 3 | | | | | | | | | | | | Year 4 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | M1 | M2 | M3 | M4 | M5 | M6 | M7 | M8 | M9 | M10 | M11 | M12 | M13 | M14 | M15 | M16 | M17 | M18 | M19 | M20 | M21 | M22 | M23 | M24 | M25 | M26 | M27 | M28 | M29 | M30 | M31 | M32 | M33 | M34 | M35 | M36 | M37 | M38 | M39 | M40 | M41 | M42 |
| 5 | **LinkedTV platform** | 1 | 42 | 42 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T5.1 | Integrating platform | 1 | 24 | 24 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T5.2 | Back-end architecture specification | 1 | 24 | 12 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T5.3 | Back-end development and component integration | 7 | 36 | 30 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T5.4 | MediaCanvas API | 7 | 30 | 12 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T5.5 | Clickable Video Playback | 7 | 30 | 12 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T5.6 | Front-end development and component integration | 7 | 36 | 30 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T5.7 | Validation | 37 | 42 | 6 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 6 | **Scenarios** | 1 | 42 | 42 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T6.1 | Media and content preparation | 1 | 30 | 12 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T6.2 | Scenario realization on LinkedTV platform | 7 | 36 | 18 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T6.3 | User trials and evaluation | 13 | 42 | 18 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

| WP | WP and Task Descriptions | Start (Mo) | End (Mo) | Duration (Mo) | Year 1 | Year 2 | Year 3 | Year 4 |
|---|---|---|---|---|---|---|---|---|
| | | | | | M1 M2 M3 M4 M5 M6 M7 M8 M9 M10 M11 M12 | M13 M14 M15 M16 M17 M18 M19 M20 M21 M22 M23 M24 | M25 M26 M27 M28 M29 M30 M31 M32 M33 M34 M35 M36 | M37 M38 M39 M40 M41 M42 |
| 7 | **Dissemination** | 1 | 42 | 42 | | | | |
| T7.1 | Dissemination and standardisation planning | 1 | 6 | 6 | | | | |
| T7.2 | Project website | 1 | 2 | 2 | | | | |
| T7.3 | Production of dissemination materials and publications | 7 | 42 | 36 | | | | |
| T7.4 | Project presentations at events | 7 | 42 | 36 | | | | |
| T7.5 | Project demonstrator | 19 | 42 | 24 | | | | |
| T7.6 | Standardisation activities | 7 | 42 | 36 | | | | |
| | | | | | | | | |
| 8 | **Markets, Businnes Models and Exploitation Strategies** | 1 | 42 | 42 | | | | |
| T8.1 | Product & market surveys | 1 | 36 | 30 | | | | |
| T8.2 | Business models | 7 | 36 | 30 | | | | |
| T8.3 | Exploitation planning & Exploitation activities | 1 | 42 | 42 | | | | |

| WP | WP and Task Descriptions | Start (Mo) | End (Mo) | Dura-tion (Mo) | Year 1 | | | | | | | | | | | | Year 2 | | | | | | | | | | | | Year 3 | | | | | | | | | | | | Year 4 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | M1 | M2 | M3 | M4 | M5 | M6 | M7 | M8 | M9 | M10 | M11 | M12 | M13 | M14 | M15 | M16 | M17 | M18 | M19 | M20 | M21 | M22 | M23 | M24 | M25 | M26 | M27 | M28 | M29 | M30 | M31 | M32 | M33 | M34 | M35 | M36 | M37 | M38 | M39 | M40 | M41 | M42 |
| **9** | **Management** | 1 | 42 | 42 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T9.1 | Project Management | 1 | 42 | 42 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T9.2 | Periodic Scientific Reports | 1 | 42 | 42 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T9.3 | Periodic Financial Reports | 1 | 42 | 42 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T9.4 | Work self-assessment and workplan revision | 7 | 42 | 18 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| T9.5 | Internal communication structures | 1 | 42 | 42 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

## Work package descriptions

| Workpackage number | 1 | | Start date or starting event: | | | M1 |
|---|---|---|---|---|---|---|
| Workpackage title:  Intelligent hypervideo analysis | | | | | | |
| Activity type[20]:  RTD | | | | | | |
| Participant id | 1 | **2** | 3 | 10 | 12 | |
| Person-months per participant: | 36 | **62** | 21 | 28 | 3 | |

**Objectives**

- *Automatically and semi-automatically decompose the audiovisual material (using segmentation, tracking), and represent and evaluate the similarity of content segments, in order to support the annotation process.*
- *Associate content segments with a first form of semantic descriptions of it, i.e. with an appropriate multitude of object and/or scene labels.*
- *Analyse text and audio information that is complementary to the visual stream, and use input from other modalities to optimize the functionality of individual annotation processes.*
- *Enrich and extend the previous analysis results, by further labeling content segment with event labels and by identifying specific instances of objects and events.*
- *Integrate all developed analysis and labeling techniques in an annotation tool for hypervideo, which will support users in the semi-automatic linking of content with minimal effort.*
- *Thoroughly evaluate the developed techniques, both internally and by participation to international benchmarking activities.*

**Description of work (broken down into tasks) and role of partners**

**T1.1 Visual information preprocessing and representation (CERTH-ITI, FRAUNHOFER, RBB) (M1-24)**
*T1.1.1: Requirements analysis and review of the state-of-the-art (M1-6)*
*T1.1.2: Content decomposition techniques for LinkedTV (M7-24)*
*T1.1.3: Content representation and similarity-based matching techniques for LinkedTV (M7-24)*

**T1.2 Visual object and scene labeling (CERTH-ITI, UEP) (M1-36)**
*T1.2.1: Requirements analysis and review of the state-of-the-art (M1-6)*
*T1.2.2: Learning and clustering techniques for LinkedTV (M7-36)*
*T1.2.3: Dimensionality reduction and GPU-based processing for computational efficiency (M7-36)*

**T1.3 Complementary text and audio analysis (S+V, CERTH-ITI, UEP, RBB) (M1-24)**
*T1.3.1: Requirements analysis and review of the state-of-the-art (M1-6)*
*T1.3.2: Components for state of the art audio and text analysis (M7-12)*
*T1.3.3: Extensions/improvements in audio/text analysis techniques for LinkedTV (M13-24)*

**T1.4 Event and instance-based labeling of visual information (CERTH-ITI, FRAUNHOFER) (M7-36)**
*T1.4.1: Requirements analysis and review of the state-of-the-art (M7-12)*
*T1.4.2: Event detection and event-based content labeling (M13-36)*
*T1.4.3: Instance-based content labeling using information from multiple modalities (M13-36)*

**T1.5 User assisted annotation tool (FRAUNHOFER, CERTH-ITI, RBB) (M7-36)**
*T1.5.1: Selection of baseline annotation tool and requirements analysis for its extension (M7-12)*
*T1.5.2: Intermediate version of the annotation tool (LinkedTV annotation tool, first release) (M13-24)*
*T1.5.3: Final version of the annotation tool (LinkedTV annotation tool, final release) (M25-36)*

---

[20] Please indicate one activity per work package:

RTD = Research and technological development (including any activities to prepare for the dissemination and/or exploitation of project results, and coordination activities); DEM = Demonstration; MGT = Management of the consortium; OTHER = Other specific activities, if applicable in this call.

**T1.6 Intelligent hypervideo analysis evaluation (CERTH-ITI, S+V, FRAUNHOFER) (M13-42)**
*T1.6.1: Internal evaluation of WP1 technologies (M13-42)*
*T1.6.2: Participation to international benchmarking activities with WP1 technologies (M13-42)*

**Deliverables (brief description) and month of delivery**
*D1.1 State of the art and requirements analysis for hypervideo (M12)*
*D1.2 Visual, text and audio information analysis for hypervideo, first release (M18)*
*D1.3 LinkedTV annotation tool, first release (M24)*
*D1.4 Visual, text and audio information analysis for hypervideo, final release (M36)*
*D1.5 LinkedTV annotation tool, final release (M36)*
*D1.6 Intelligent hypervideo analysis evaluation, final results (M42)*

**Milestones and expected result**

*(M12) Requirements analysis and review of state of the art for hypervideo, and selection of baseline annotation tool*
*(M24) Implementation of LinkedTV analysis techniques and hypervideo annotation tool, first release*
*(M36) Implementation of LinkedTV analysis techniques and hypervideo annotation tool, final release*
*(M42) Final evaluation results*

| Workpackage number | 2 | Start date or starting event: | | | M1 |
|---|---|---|---|---|---|
| Workpackage title:  Linking hypervideo to web content | | | | | |
| Activity type[21]:  RTD | | | | | |
| Participant id | 2 | 3 | **5** | 8 | |
| Person-months per participant: | 12 | 52 | **59** | 18 | |

**Objectives**

- *Specify and implement a URI-based mechanism for addressing media fragments.*
- *Design lightweight annotation schemas for media objects, interoperable with existing multimedia standards.*
- *Research web mining techniques on semi-structured and unstructured data to annotate media objects with concepts and expand the Web-based metadata layer to become a Connected Media Layer.*
- *Develop the infrastructure to crawl relevant parts of the Web and mine concepts, creating and searching the Connected Media Layer.*

**Description of work (broken down into tasks) and role of partners**

**T2.1 Media fragments addressing (EURECOM, CERTH-ITI)(M1-36)**

*T2.1.1 Specification for media fragment addressing (M1-24)*

This sub-task will provide a URI-based mechanism for addressing fragments of multimedia content together with client-side and server-side implementations passing test-cases.

*T2.1.2 Extension of Web search & retrieval to support media fragments (M13-36)*

This sub-task will enhance the search and retrieval modules with the media fragments specification.

**T2.2 Lightweight metadata models for hypervideo (EURECOM, UEP, CERTH-ITI)(M1-36)**

*T2.2.1 Design of a lightweight metadata model and alignment with other schemas (M1-18)*

This sub-task will model a set of lightweight metadata ontologies for multimedia, compatible with existing standards.

*T2.2.2 API implementation and refinements of the metadata models (M13-36)*

This sub-task will implement, test and publicly release an API of the metadata models.

**T2.3 Web mining for support of media annotation (UEP, EURECOM)(M1-42)**

*T2.3.1 Selection and evaluation of methods for Web mining (M1-6)*

This sub-task will evaluate techniques and tools (e.g. OpenCalais, GATE,
SPROUT) for extracting information from web resources in the context of
media analysis.

*T2.3.2 Specification of Web mining process for improved concept identification in hypervideo (M7-36)*

This sub-task will research, develop and evaluate novel mining techniques for identifying and disambiguating concepts in hypervideo.

*T2.3.3 Integration and refinement of concept labelling through Web content mining process (M13-42)*

This sub-task will complement specifically the multimedia analysis performed in WP1 for labelling concepts in hypervideo.

**T2.4 Retrieving additional content from the web  (UEP, CONDAT, EURECOM)(M13-42)**

*T2.4.1 Specification of the Linked Media Layer (M13-24)*

---

[21] Please indicate one activity per work package:

RTD = Research and technological development (including any activities to prepare for the dissemination and/or exploitation of project results, and coordination activities); DEM = Demonstration; MGT = Management of the consortium; OTHER = Other specific activities, if applicable in this call.

This sub-task will specify the Linked Media Layer composed of a fine-grained mechanism for addressing fragments of multimedia and annotations schemas.

*T2.4.2 Search and Retrieval in the Linked Media Layer (M25-42)*

This sub-task will iteratively integrate the media fragment specification and the metadata schemas together with the annotations gathered from web mining processes in order to provide a search and retrieval module in the Linked Media Layer.

*T2.4.2 Web mining approaches to annotating Web-based content and gathering information around concepts (M13-42)*

This sub-task will research and evaluate additional Web mining techniques on completely unstructured data to enrich the Linked Media Layer

---

**Deliverables (brief description) and month of delivery**

*D2.1 Specification of the Media Fragment URI scheme (M6)*
*D2.2 Specification of lightweight metadata models for multimedia annotation (M9)*
*D2.3 Specification of Web mining process for hypervideo concept identification (M12)*
*D2.4 Annotation and retrieval module of media fragments (M24)*
*D2.5 Specification of the Linked Media Layer (M24)*
*D2.6 Advanced concept labelling by complementary Web mining (M36)*
*D2.7 Final Linked Media Layer and evaluation (M42)*

---

**Milestones and expected result**

*(M12) Media fragment addressing and multimedia metadata (first specification), Web mining for concept identification and disambiguation (first specification)*
*(M24 ) Media fragment test cases and implementation, Web mining for concept annotation and specification of the Linked Media Layer*
*(M36 ) Multimedia metadata API (final specification), Web mining for concept identification and disambiguation (final specification)*
*(M42 ) Evaluation and final results*

| Workpackage number | 3 | Start date or starting event: | | | M1 |
|---|---|---|---|---|---|
| **Workpackage title: LinkedTV interface and presentation engine** | | | | | |
| **Activity type[22]: RTD** | | | | | |
| **Participant id** | **4** | 6 | 9 | 10 | 11 | 12 |
| **Person-months per participant:** | **58** | 5 | 21 | 5 | 9 | 6 |

**Objectives**

- *Collect user interface and system requirements for information interfaces*
- *Design a number of cross-scenario information interfaces*
- *Specify a flexible and adaptable description scheme for interfaces*
- *Implement information selection and browsing engine for concept-based presentations*
- *Compiling design guidelines for concept-based information interfaces*

**Description of work (broken down into tasks) and role of partners**

T3.1 **User requirements analysis** (USG, S+V, UMONS, RBB)(M1-6)
*T3.1.1 Identify a number of cross-scenario information selection/browsing tasks (M1-3)*
This sub-task will identify common user information needs based on the three scenarios envisaged in WP6.
*T3.1.2 Establish user interface and system functionality requirements (M4-6)*
This sub-task will specify the functionality requirements based on the previous analysis.

T3.2 **User-centred design and specification of hypervideo information interfaces** (CWI, UMONS, RBB)(M1-12, M25-36)
*T3.2.1 Design of first user interfaces (M1-6)*
This subtask will design two or three mockup interfaces per scenario based on the requirements previously defined.
*T3.2.2 Development of first user interfaces (M7-12)*
This sub-task will sketch the interfaces previously designed.
*T3.2.3 Selection of final user interface design and specification (M25-36)*
After an evaluation phase, the final interfaces for each scenario will be chosen among the various sketch and further specified and developed.

T3.3 **Develop the interface and presentation engine** (NOTERIK, CWI) (M7-36)
*T3.3.1 Development of the interface and presentation engine (M7-12)*
This sub-task will develop a first version of the presentation engine, part on the LinkedTV platform.
*T3.3.2 Refinement of the interface and presentation engine (M13-36)*
This sub-task will develop the final presentation engine, after an evaluation phase.

T3.4 **Evaluate functionalities and interfaces of the connected media and background information** (CWI, S+V, RBB, UMONS)(M19-24, M37-42)
*T3.4.1 First evaluation (M19-24)*
This sub-task will conduct the first user evaluation of the various interfaces developed
*T3.4.2 Final evaluation (M37-42)*
This sub-task will conduct the second user evaluation of the final interfaces chosen and the overall presentation engine.

**Deliverables (brief description) and month of delivery**

*D3.1 Specification of functionality requirements satisfying user information needs (M6)*
*D3.2 Specification of presentation interfaces for the three scenarios (M6)*

---

[22] Please indicate one activity per work package:

RTD = Research and technological development (including any activities to prepare for the dissemination and/or exploitation of project results, and coordination activities); DEM = Demonstration; MGT = Management of the consortium; OTHER = Other specific activities, if applicable in this call.

*D3.3 LinkedTV user interfaces sketch (M12)*
*D3.4 LinkedTV interface and presentation engine version 1 (M12)*
*D3.5 Requirements document for LinkedTV user interfaces (version 2) (M24)*
*D3.6 LinkedTV interface and presentation engine version 2 (M24)*
*D3.7 LinkedTV user interfaces selected and refined (version 2) (M36)*
*D3.8 Design guideline document for concept-based presentations (M42)*

**Milestones and expected result**

*M3  Initial descriptions of scenario tasks as input for discussions with scenario partners.*
*M6  User groups identified with which to try out sketch and prototype interfaces.*
*M9  Initial sketches and mockups of interfaces available for discussions with scenario partners.*
*M12 First interfaces created for testing with scenario partners,first presentation engine ready for LinkedTV platform.*
*M24  Refined presentation engine.*
*M36  Final interfaces decided and implemented.*
*M42  Final evaluation of interfaces in the user trials.*

| Workpackage number | 4 | | | Start date or starting event: | | | M1 | |
|---|---|---|---|---|---|---|---|---|
| **Workpackage title: Contextualisation and personalisation** | | | | | | | | |
| **Activity type[23]: RTD** | | | | | | | | |
| **Participant id** | **1** | 2 | 3 | 4 | 7 | 11 | | |
| **Person-months per participant:** | **36** | 31 | 20 | 8 | 12 | 24 | | |

**Objectives**
- *Specify a schema for user profiles*
- *Specify a schema for contextualised profile and knowledge pull*
- *Specify a means to filter and select concepts and content on the basis of user profiles*
- *Extend and refine user profile on the basis of user activity and social profiles*

**Description of work (broken down into tasks) and role of partners**

**T4.1 User profiling (CERTH-ITI, FRAUNHOFER, UEP, CWI, STI) (M1-M24)**

*T4.1.1 State of the art and requirements analysis for user profiling (M1-6)*

This subtask will provide a study of the user requirements and a comprehensive research and evaluation of the state-of-the-art approaches on user profiling for multimedia content.

*T4.1.2 First specification of user profile schema and processes for profile capturing (M7-12)*

Specification of the schema for semantic representation of the user profile and mechanism for unobtrusive capturing and updating of user preferences through his/her transactions.

*T4.1.3 Extensions/refinements to the user profile schema and processes for profile capturing (M13-18)*

Refined implementation for semantically representing of the user profile and advanced mechanism for profile learning and understanding.

*T4.1.4 Final specification of user profile schema and processes for profile capturing (M19-24)*

Final implementation and testing of the user profile representation and learning mechanisms.

**T4.2 User Behaviour and Contextualisation (UMONS, FRAUNHOFER, CERTH-ITI) (M1-M36)**

*T4.2.1 State of the art and requirements analysis for contextualisation (M1-6)*

This subtask will provide a study of the current research and evaluation of the state-of-the-art approaches on user behaviour tracking

*T4.2.2 First specification for personalisation based on user behaviour (M7-12)*
Mechanism for identifying user interests based on their physical behaviour, e.g. eye tracking allows to identify which concepts in video catch and hold user attention.

*T4.2.3 Extensions/refinements for personalisation based on user behaviour (M13-24)*
Mechanism for identifying and extracting the semantic knowledge in user behavioural tracking.

*T4.2.4 Advanced concept-based analysis for personalisation based on user behaviour (M25-36)*
This subtask will provide the mechanism to expand the personalisation of the LinkedTV services with the aid of external web resources based on available high-level multimedia concepts for filtering and contextualising the identified concepts and content of interest to the user.

**T4.3 Personalized content filter (FRAUNHOFER, CERTH-ITI, UEP, STI) (M7-M36)**

---

[23] Please indicate one activity per work package:
RTD = Research and technological development (including any activities to prepare for the dissemination and/or exploitation of project results, and coordination activities); DEM = Demonstration; MGT = Management of the consortium; OTHER = Other specific activities, if applicable in this call.

*T4.3.1 Development of first profile matcher and content filter (M7-12)*
First implementation of the profile-content semantic matcher, where provided concepts and content will be filtered based on the user profile.

*T4.3.2 Refined profile matcher and content filter (M19-24)*
Refined implementation of the profile-content semantic matcher, integrating advanced knowledge pulling.

*T4.3.3 Final profile matcher and content filter (M31-36)*
Final implementation of the profile-content semantic matcher, incorporating advanced concept and content filtering.

**T4.4 Evaluation of personalization technologies (FRAUNHOFER, UMONS, CERTH-ITI) (M37-M42)**
This task will define the user studies, test beds and evaluation metrics for testing the personalization module and provide an experimental evaluation of the module.

---

**Deliverables (brief description) and month of delivery**

*D4.1 Specification of user profiling and contextualisation (M6)*
*D4.2 User profile schema and profile capturing (M12)*
*D4.3 Content and concept filter v1 (M12)*
*D4.4 User profile and contextual adaptation (M24)*
*D4.5 Content and concept filter v2 (M24)*
*D4.6 Contextualisation solution and implementation (M36)*
*D4.6 Evaluation and final results (M42)*

---

**Milestones and expected result**

*(M12) User profile specification, processes for user profile capture and content filtering/adaptation,initial personalization filter component*
*(M24)   Final user profile specification, processes for user profile capture and refined/extended content filtering/adaptation, revised personalization filter component*
*(M36) Final specifications for content filtering/adaptation, final personalization filter component*

*(M42) Final results and evaluation*

| Workpackage number | 5 | | | Start date or starting event: | | M1 | |
|---|---|---|---|---|---|---|---|
| Workpackage title: LinkedTV platform | | | | | | | |
| Activity type[24]: RTD | | | | | | | |
| Participant id | 1 | 2 | 4 | 5 | | **8** | 9 |
| Person-months per participant: | 6 | 4 | 4 | 4 | | **48** | 21 |

**Objectives**

*This WP realises the integration of components of LinkedTV into an end-to-end platform. It will integrate or provide the following components on the back-end:*
- *audio / video and Web resource analysis and annotation*
- *local (meta)data repository and query functionality*
- *search and retrieval of media from the Web (Connected Media Layer)*
- *linking of all networked multimedia materials*
- *web crawler for periodic link update*
- *caching for efficient user quality of experience*
- *platform management (e.g. component monitoring)*
- *output of video, associated content and metadata to network delivery servers such as Web or IPTV*

*It will integrate or provide the following components on the front-end:*
- *clickable video playback*
- *interactive content layer for the LinkedTV user interfaces (the MediaCanvas API)*
- *user profile capture and personalization filter*
- *front-end cache for related content*

**Description of work (broken down into tasks) and role of partners**

**T5.1 Integrating platform (CONDAT, NOTERIK)(M1-24)**
*T5.1.1 Specification and set-up of the integrating platform (M1-6)*
*T5.1.2 Development of delivery modules for Web TV and IPTV (M7-12)*
*T5.1.3 Testing of delivery over Web TV and IPTV (M13-18)*
*T5.1.4 Integration platform testing and reconfiguration (M19-24)*

**T5.2 Back-end architecture specification (CONDAT, NOTERIK)(M1-6, 19-24)**
*T5.2.1 First LinkedTV architecture specification (M1-6)*
*T5.2.2 Architectural validation and refinement (M19-24)*

**T5.3 Back-end development and component integration (CONDAT, NOTERIK)(M7-36)**
*T5.3.1 Initial set-up of the back-end system and specification of component integration (M7-12)*
*T5.3.2 Integration of components onto the back-end system (M13-18)*
*T5.3.3 Refinement of back-end system (M19-24)*
*T5.3.4 Integration of new components onto the back-end system (M25-30)*
*T5.3.5 Final refinement of back-end system (M31-36)*

**T5.4 MediaCanvas API (NOTERIK, CWI)(M7-12, 25-30)**
*T5.4.1 First MediaCanvas API and display tool (M7-12)*
*T5.4.2 Second MediaCanvas API and display tool (M25-30)*

**T5.5 Clickable Video Playback (NOTERIK, EURECOM)(M7-12, 25-30)**
*T5.5.1 First Clickable Video player (M7-12)*
*T5.5.2 Second Clickable Video player (M25-30)*

---

[24] Please indicate one activity per work package:
RTD = Research and technological development (including any activities to prepare for the dissemination and/or exploitation of project results, and coordination activities); DEM = Demonstration; MGT = Management of the consortium; OTHER = Other specific activities, if applicable in this call.

**T5.6 Front-end development and component integration (NOTERIK, CERTH-ITI)(M7-36)**
*T5.6.1 Initial set-up of the front-end system and specification of component integration (M7-12)*
*T5.6.2 Integration of components onto the front-end system (M13-18)*
*T5.6.3 Refinement of front-end system (M19-30)*
*T5.6.4 Final refinement of back-end system (M31-36)*

**T5.7 Validation (FRAUNHOFER, CONDAT, NOTERIK)(M37-42)**
*T5.7.1 Validation of the front-end (M37-42)*
*T5.7.2 Validation of the back-end (M37-42)*
*T5.7.3 Validation of the end-to-end platform (M37-42)*

---

**Deliverables (brief description) and month of delivery**
*D5.1 LinkedTV platform and architecture (M6)*
*D5.2 LinkedTV front-end: video player and MediaCanvas API (M12)*
*D5.3 First LinkedTV end-to-end platform (M18)*
*D5.4 Final LinkedTV integrating platform (M24)*
*D5.5 LinkedTV front-end: video player and MediaCanvas API v2 (M30)*
*D5.6 Final LinkedTV end-to-end platform (M36)*
*D5.7 Validation of the LinkedTV architecture (M42)*

---

**Milestones and expected result**

*(M6) First specifications of the LinkedTV architecture and integrating platform*
*(M12) Development/deployment of the WP5 specific back-end and front-end components, specification for the integration of the components from WPs 1-4*
*(M18) First end-to-end platform ready with integrated back-end and front-end systems and functioning delivery modules for Web TV and IPTV*
*(M24)  Final specification of the integrating platform*
*(M30) Final development/deployment of components in the back-end and front-end*
*(M36) Final implementation of the end-to-end platform*
*(M42) Validation of the end-to-end platform*

| Workpackage number | 6 | | | **Start date or starting event:** | | | | M1 | |
|---|---|---|---|---|---|---|---|---|---|
| **Workpackage title: Scenarios** | | | | | | | | | |
| **Activity type[25]:  RTD** | | | | | | | | | |
| **Participant id** | 1 | 2 | 3 | 5 | 6 | **7** | 10 | 11 | 12 | |
| **Person-months per participant:** | 2 | 2 | 4 | 18 | 8 | **16** | 21 | 21 | 22 | |

**Objectives**

- *Define the content and concepts relevant for each scenario*
- *Storyboarding, feasibility study and realizing the scenarios on the LinkedTV platform*
- *User trials and evaluation*

**Description of work (broken down into tasks) and role of partners**

**T6.1  Media and content preparation (STI, RBB, UMONS, S+V, CERTH-ITI, UEP, EURECOM)**
*T6.1.1 Initial media and content preparation (M1-6)*
*T6.1.2 Revised media and content preparation (M25-30)*

**T6.2  Scenario realization on LinkedTV platform (STI, RBB, UMONS, S+V, FRAUNHOFER)**
*T6.2.1 Storyboarding of the scenarios (M7-12)*
*T6.2.2 Scenario realization on the LinkedTV platform (M19-24)*
*T6.2.3 Revised scenario realization on the LinkedTV platform (M31-36)*

**T6.3  User trials and evaluation (UNISG, EURECOM, RBB, UMONS, S+V, STI)**
*T6.3.1 Internal trials with the partners (M13-18)*
*T6.3.2 User trials held by the partners (M25-30)*
*T6.3.3 Final user trials held by the partners (M37-42)*

**Deliverables (brief description) and month of delivery**
*D6.1 Scenario descriptions (M12)*
*D6.2 Scenario demonstrators (M24)*
*D6.3 User trial results (M30)*
*D6.4 Scenario demonstrators (M36)*
*D6.5 Final evaluation (M42)*

**Milestones and expected result**

*(M6)  The media and content required has been selected*
*(M12)  Scenarios have been described*
*(M18) Scenario preparation is complete and has been trialed internally among the partners*
*(M24)  Scenarios have been realized on LinkedTV platform and demonstrators are available*
*(M30) User trials have taken place arranged by the partners*
*(M36) Scenarios have been refined and realized anew on the LinkedTV platform*
*(M42) Final user trials have taken place and the scenarios have been evaluated*

---

[25] Please indicate one activity per work package:

RTD = Research and technological development (including any activities to prepare for the dissemination and/or exploitation of project results, and coordination activities); DEM = Demonstration; MGT = Management of the consortium; OTHER = Other specific activities, if applicable in this call.

| Workpackage number | 7 | Start date or starting event: | | | | | | M1 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Workpackage title: Dissemination | | | | | | | | | | | | |
| Activity type[26]:  OTHER | | | | | | | | | | | | |
| Participant id | 1 | 2 | 3 | 4 | 5 | 6 | **7** | 8 | 9 | 10 | 11 | 12 |
| Person-months per participant: | 6 | 3 | 3 | 3 | 3 | 3 | **20** | 2 | 2 | 2 | 6 | 7 |

**Objectives**

*The aim of this workpackage is to ensure the dissemination of information about the project, its objectives, approaches and results. This will be achieved through a public website, dissemination materials, web based dissemination, participation and publication at academic and industry events as well as in scientific and trade journals, and the preparation of demonstrators which will showcase LinkedTV services, both at events and remotely (over Web channels). We will also push for standardisation of project results, particularly with respect to a URI-based mechanism for addressing spatio-temporal multimedia fragment within the W3C and an ontology and its API for representing multimedia metadata within W3C, EBU and IPTC.*

**Description of work (broken down into tasks) and role of partners**

**T7.1 Dissemination and standardisation planning (STI)(M1-6)**

**T7.2 Project website (STI)(M1-2)**

**T7.3 Production of dissemination materials and publications (STI)(M7-42)**

**T7.4 Project presentations at events (STI)(M7-42)**

**T7.5 Project demonstrator (STI)(M19-42)**
*T7.5.1 First project demonstrator (M19-24)*
*T7.5.2 Dissemination of the demonstrator online and at events (M25-30)*
*T7.5.3 Second project demonstrator (M31-36)*
*T7.5.4 Dissemination of the demonstrator online and at events (M37-42)*

**T7.6 Standardisation activities (STI,EURECOM)(M7-42)**
*T7.6.1 First standardisation push (M7-18)*
*T7.6.2 Evaluation of standardisation efforts (M19-24)*
*T7.6.3 Second standardisation push (M25-36)*
*T7.6.4 Sustainability actions for standardisation activities after the project (M37-42)*

**Deliverables (brief description) and month of delivery**
*D7.1 Project website (M2)*
*D7.2 Dissemination and standardisation plan (M6)*
*D7.3 Dissemination  and standardisation report v1 (M18)*
*D7.4 Project demonstrator v1 (M24)*
*D7.5 Dissemination and standardisation report v2 (M30)*
*D7.6 Project demonstrator v2 (M36)*
*D7.7 Dissemination and standardisation report v3 (M42)*

**Milestones and expected result**
*(M6) Plan for dissemination and standardisation during the project is fixed, project website is online and*

---

[26] Please indicate one activity per work package:

RTD = Research and technological development (including any activities to prepare for the dissemination and/or exploitation of project results, and coordination activities); DEM = Demonstration; MGT = Management of the consortium; OTHER = Other specific activities, if applicable in this call.

*publicly accessible*
*(M18) First dissemination activities have taken place, first push for standardisation of project results*
*(M24) First demonstrators for the project are ready*
*(M30) First use of the demonstrators at academic, industrial or public events*
*(M36) A final demonstrator is ready for presenting the results of the project , another push for standardisation of project results*
*(M42) Final actions disseminating and demonstrating the project results have taken place, sustainability actions are in place to ensure project results are standardised after the project ends*

| Workpackage number | 8 | | Start date or starting event: | | | | | M1 | |
|---|---|---|---|---|---|---|---|---|---|
| **Workpackage title:  Markets, Business Models and Exploitation Strategies** | | | | | | | | | |
| **Activity type[27]:  RTD** | | | | | | | | | |
| **Participant id** | | 1 | **6** | 7 | 8 | 9 | 10 | 11 | 12 | |
| **Person-months per participant:** | | 10 | **30** | 10 | 6 | 6 | 6 | 6 | 3 | |

**Objectives**
- *To provide a detailed market overview of the current state-of-the art market developments and future trends based on a broad market analysis*
- *Development of innovative business models for different players involved in the LinkedTV value chain*
- *To develop an exploitation plan as a guideline for exploitation activities during the project as well as individual and common exploitation strategies that will assure that LinkedTV results can be turned in commercial success*

**Description of work (broken down into tasks) and role of partners**

**T8.1 Product & Market Surveys (UNISG, S+V, UMONS, RBB, CONDAT, NOTERIK, STI)**

*T8.1.1 First market survey of user needs (M1-6)*

The goal of this first market analysis is to explore how IPTV is used by users today and what their preferences for interactive IPTV would be. The focus of the analysis will be on preferences of users for different innovative interactive content formats. The analysis will be based on the one hand on secondary information sources, and on the other hand on primary quantitative and qualitative market research.

*T8.1.2 First market survey of content and provider needs (M7-12)*

The aim of the market analysis of content providers and providers is to analyse their requirements. Important questions that will be considered are preferences for content that needs to be connected, potential interaction content format, existing experiences and in particular how new interactive content formats may fit into existing routines. The analysis will be based on secondary analysis and comparison with similar formats in online media. In addition, the content providers of the project will be consulted together with external providers and content providers.

*T8.1.3 Second market survey of user needs (M19-24)*

The second market analysis will take place when the first version of the LinkedTV technology is designed. The concrete concepts will be evaluated by the user. The goaö will be to evaluate the potential adoption of LinkedTV services. At the same time the data resulting from quantitative and qualitative market research will be compared with the data from the first market analysis. In addition to that the findings based on secondary data will be updated.

*T8.1.4 Second market survey of content and provider needs (M25-30)*

The second market analysis will evaluate the acceptance of the proposed LinkedTV solution by content providers. In addition the findings from the secondary sources and the primary market research will be compared. The emphasize will furthermore be on potential new advertising models and their acceptance by content providers.

*T8.1.5 Overall market analysis for exploitation of LinkedTV results (M31-36)*

The final market analysis will be an update on the findings from secondary and primary market research. It will be input for the final version of the exploitation plans.

**T8.2 Business Models (UNISG, FRAUNHOFER, STI)**

---

[27] Please indicate one activity per work package:

RTD = Research and technological development (including any activities to prepare for the dissemination and/or exploitation of project results, and coordination activities); DEM = Demonstration; MGT = Management of the consortium; OTHER = Other specific activities, if applicable in this call.

*T8.2.1 Business model development (M7-18, 25-36)*

Business model development will be based on the results of the first market analysis, The main components of business models will be conceptualized. This are potential customers, participants in the value chain, product design in terms of content formats and user interaction, The main emphasize will be furthermore on payment models and financial flows. The business models will be developed for all players in the value chain and from perspective of advertisers. The analysis will also evaluate relevant legal aspects.

*T8.2.2 Analysis of the LinkedTV end-to-end platform value chain (M19-24)*

The second part of the business model analysis will concentrate on the analysis of the end-to-end value chain of the Connect Me platform and the relationships among involved players.

**T8.3 Exploitation Planning & Exploitation Activities (UNISG, CONDAT, NOTERIK, STI)**

*T8.3.1 Exploitation plan for during the project (M1-6)*

From the very beginning of the project exploitation activities will be considered and scheduled. The first exploitation plan will collect and summarize exploitation activities during the project. The exploitation plan will summarize all planned exploitation activities.

*T8.3.2 Exploitation activities (M7-18, 25-36)*

During the project the agreed upon exploitation activities will be carried out. This in particular includes also the development of concepts for and negotiation of joined exploitation and

*T8.3.3 Evaluation of exploitation impact (M19-24)*

After first exploitation activities are carried out, their impact will be evaluated and corrective measures will be proposed in case necessary.

*T8.3.4 Exploitation plan for after the project (M37-42)*

The ideas for common and individual exploitation plans will be collected and summarized. The experiences of exploitation activities during the project will be the used as input.

---

**Deliverables**

*D8.1 Exploitation plan for the project (M6)*

*D8.2 First market analysis (M12)*

*D8.3 LinkedTV business models, v1 (M18)*

*D8.3 Evaluation of exploitation impact and recommendations (M24)*

*D8.4 LinkedTV end-to-end platform value chain (M24)*

*D8.5 Market and product survey for LinkedTV services and technology (M36)*

*D8.6 LinkedTV business models, v2 (M36)*

*D8.7 Common and individual exploitation plans for after the project (M42)*

---

**Milestones and expected result**

*(M6) exploitation plan for the project is ready*

*(M12)– initial market analysis has been completed*

*(M18)  first business plans are prepared, first exploitation activities have taken place*

*(M24)  analysis of the LinkedTV value chain has been completed*

*(M36)  final market analysis, final business plans and final exploitation activities are complete*

*(M42)  common and individual exploitation plans for LinkedTV results have been produced*

| Workpackage number | 9 | Start date or starting event: | | | | M1 | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Workpackage title: Management** | | | | | | | | | | | | |
| **Activity type[28]: MGMT** | | | | | | | | | | | | |
| **Participant id** | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| **Person-months per participant:** | **21** | 1 | 1 | 1 | 0 | 1 | 2 | 1 | 1 | 1 | 1 | 1 |

**Objectives**

*The management workpackage covers all the activities of the management structure of the project, which have the goal of ensuring a correct and successful running of the workplan. Individual tasks and reports are intended to ensure regular reporting which can aid the co-ordinators in identifying potential risks or problems, as well as ensuring that the project is operating correctly and on track in its scientific and technological goals.*

**Description of work (broken down into tasks) and role of partners**

T9.1 Project Management (M1-42) (**FRAUNHOFER**)
T9.2 Periodic Scientific Reports (M1-42) (**FRAUNHOFER**)
T9.3 Periodic Financial Reports (M1-42) (**FRAUNHOFER**)
T9.4 Work Self-assessment and Workplan Revision (M7-12, 19-24, 37-42) (**FRAUNHOFER**)
T9.5 Internal Communication Structures: mailing lists, wiki etc. (M1-42) (**STI**)

**Deliverables (brief description) and month of delivery**
*D9.1 Annual Project Scientific Report (M12,24,36)*
*D9.2 Annual Project Financial Report (M12,24,36)*
*D9.3 Self-Assessment of Work Performed and Recommendations for Workplan Revision (M12,24)*
*D9.4 Final Project Report (M42)*

**Milestones and expected result**
*(M12, M24, M36) Annual reports on scientific and financial status of the work*
*(M12, M24) Self-assessment of work done in the project and , if necessary, recommendations for changes to the workplan*
*(M42) Final report with summary of project results and sustainability actions*

---

[28] Please indicate one activity per work package:
RTD = Research and technological development (including any activities to prepare for the dissemination and/or exploitation of project results, and coordination activities); DEM = Demonstration; MGT = Management of the consortium; OTHER = Other specific activities, if applicable in this call.

## Summary of staff effort

**Table 6: Staff Effort**

| Partic. no. | Partic. short name | WP1 | WP2 | WP3 | WP4 | WP5 | WP6 | WP7 | WP8 | WP9 | Total person months |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | *FRAUN-HOFER* | *36* | | | *36* | *6* | *2* | *6* | *10* | *21* | *117* |
| 2 | *CERTH* | *62* | *12* | | *31* | *4* | *2* | *3* | | *1* | *115* |
| 3 | *UEP* | *21* | *52* | | *20* | | *4* | *3* | | *1* | *101* |
| 4 | *CWI* | | | *58* | *8* | *4* | | *3* | | *1* | *74* |
| 5 | *EURO* | | *59* | | | *4* | *18* | *3* | | *0* | *84* |
| 6 | *USG* | | | *5* | | | *8* | *3* | *30* | *1* | *47* |
| 7 | *STI* | | | | *12* | | *16* | *20* | *10* | *2* | *60* |
| 8 | *CONDAT* | | *18* | | | *48* | | *2* | *6* | *1* | *75* |
| 9 | *NOTERIK* | | | *21* | | *21* | | *2* | *6* | *1* | *51* |
| 10 | *S+V* | *28* | | *5* | | | *21* | *2* | *6* | *1* | *63* |
| 11 | *UMONS* | | | *9* | *24* | | *21* | *6* | *6* | *1* | *67* |
| 12 | *RBB* | *3* | | *6* | | | *22* | *7* | *3* | *1* | *42* |
| **Total** | | *150* | *141* | *104* | *131* | *87* | *114* | *60* | *77* | *32* | *896* |

## List of milestones

The project has a duration of 42 months, which will split into 5 milestones. These milestones represent a cycle of research, development as part of an overall architecture and evaluation through the three scenarios. This is illustrated in Figure 16.
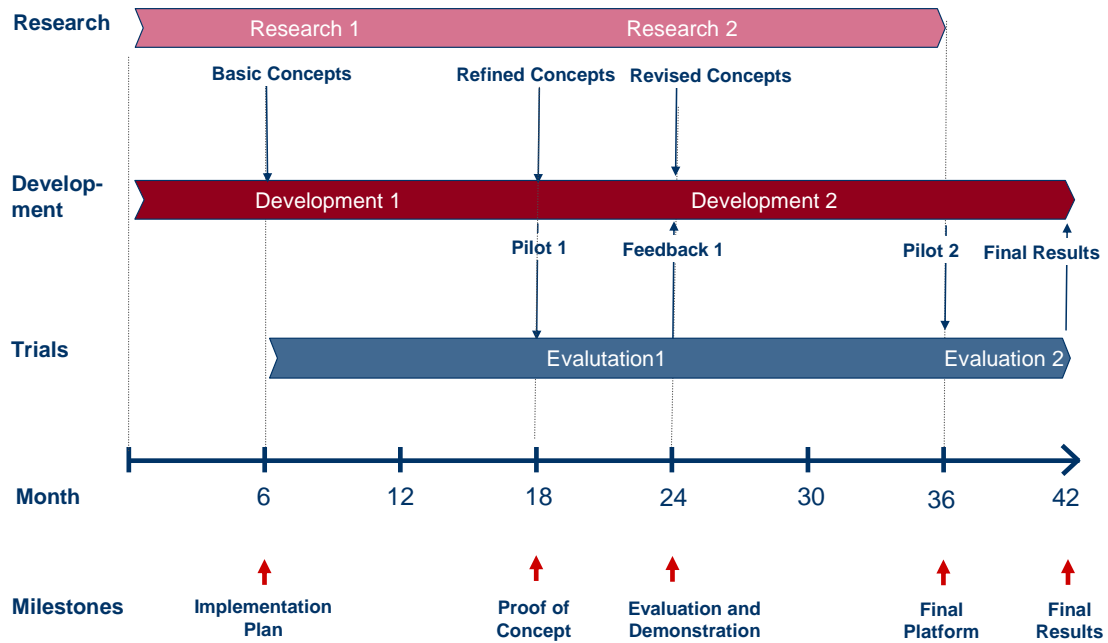
**Figure 16: Project milestones**

# Section 2.  Implementation

## *2.1. Management structure and procedures*

A clear management structure and well defined procedures are vital for the overall success of the project, independent from scientific and technological expertise or a high potential impact. Hence, in LinkedTV, we take care to define the project management and to ensure the resources and commitment that are needed to enforce the management structure and ensure a correct management of all activities of the project.

This management structure will be formally agreed upon by all members of the consortium. The presented workplan will form the basis for management decisions, e.g. in the assessment of project performance the stated milestones, task aims and development stages will be used. Management will also adhere to the agreed responsibilities and rights of each partner and will be endued with the necessary powers to enforce compliance to the workplan and to management decisions by the consortium. Furthermore, all this will happen in compliance with Ec regulations and practices.

## 2.1.1. Project workplan structure

Project work is divided into workpackages, each focused on a particular aim, and with a stated plan to achieve that aim. They are split into individual tasks, deliverables as a means of reporting on achievement of tasks, and milestones as a means of measuring at predefined intervals the achievement of planned tasks. Furthermore, we have explicitly identified dependencies between workpackages which must be taken into consideration, e.g. that a delay in work performed in one workpackage may delay other work. Workpackages each have a workpackage leader who is responsible for the overall activity of the workpackage. Tasks also have allocated leaders who take responsibility for the completion of individual tasks, including the deliverable(s) to be produced as a result of their task. These leaders report to the workpackage leader on activities in their tasks and deliverables. All leaders co-ordinate the participating partners in the achievement of the work as specified by the overall workplan.

## 2.1.2. Management structure

The management structure of LinkedTV is illustrated in Figure 10. The **project coordinator** will have overall responsibility for the project. **Dr. Joachim Kohler (Fraunhofer)** is an experienced researcher and project coordinator who provides all necessary capabilities for this role. He has been technical co-ordinator of the European IP-project LIVE as well as deeply active in the CHORUS coordination action on multimedia search. Together with the management board he will take all necessary decisions to lead the project. He will coordinate the work of the technical board with all work package leaders including the needed communication between them. The communication to the European Commission will be one of his main responsibilities – including annual review meetings, regular management reports, dissemination actions at European level, etc. The performance of partners will be checked at regular intervals by the project co-ordinator. He will organize full project meetings on a regular basis. He will supervise the management office which takes care of the budget, the quality management, risk management, and IPR management. The project coordinator will organize the activities of the advisory board and its communication with the project.

LinkedTV as an Integrated Project will be a complex endeavour with many related efforts. In order to manage this complex research the project coordinator will be supported by

a **scientific director**. He takes charge of ensuring the scientific and technological contributions of the work and of the scientific coherence of all partners' contributions in the work packages and tasks. The project's scientific director will be **Dr. Lyndon Nixon (STI International)**. The scientific director will pay special attention to the activities in the technical board and together with the project coordinator guarantee the necessary coordination between them.

The **technical board** will be responsible for the coordination of all work packages and tasks. It is lead by the project coordinator who is supported by the scientific director. All work package leaders work together in this board. Decisions related to the scientific and technological work of the project will be made on a majority basis. An assessment of the project's performance will be made at regular intervals in meetings of the project's technical board.

The **project office** will be based at Fraunhofer. This guarantees the necessary tight collaboration of the project coordinator and those people supporting him with budget, risk, quality, and other related management issues of the project.

**An advisory board** will be established made up of external experts from industry, broad casting organizations, and research institutions with a scientific interest in the project and comprehensive experiences in the project's areas.
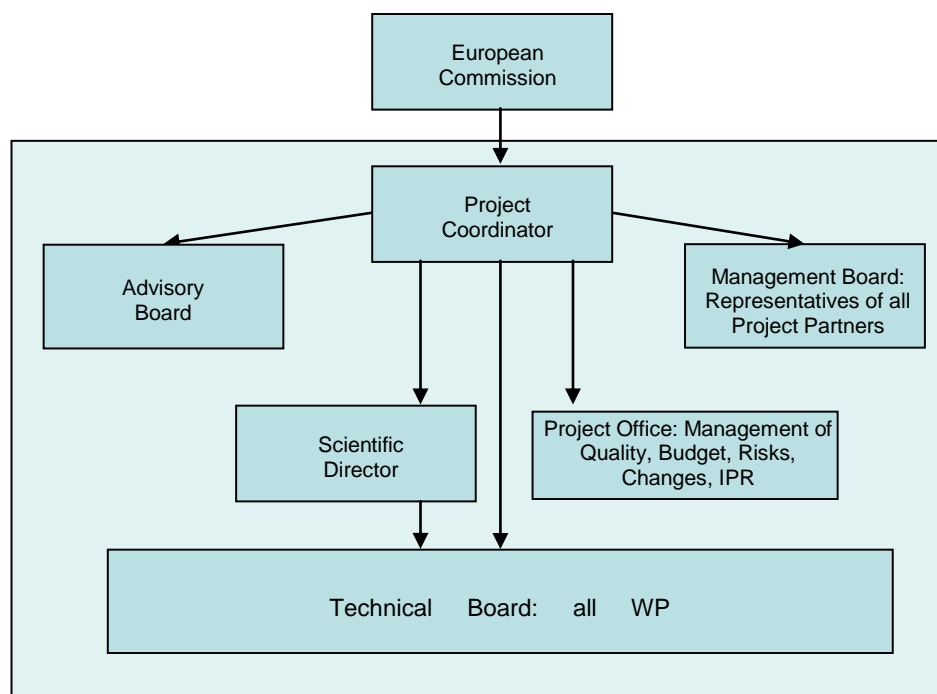
**Figure 17: LinkedTV Management Structure**

## 2.1.3. Meetings and communication

Partners in LinkedTV will meet face to face at least 4 times a year, excluding the annual review meeting. In-between, they will hold regular telephone conferences and bilateral meetings as required. Furthermore, daily activities in the project will be enabled by communication structures which also allow for monitoring and assessment.

LinkedTV will hold project meetings quarterly, in which all partners are expected to attend with adequate personnel who are capable of reporting about activities performed and future plans. These meetings are the primary gatherings of LinkedTV partners and will consist of:

1. meetings of all work packages
2. meetings of the technical and management boards
3. potentially, invitation to and participation of members of the advisory board
4. additionally, organised events as judged profitable to the project, e.g. invited speaker(s), involvement of industry observers etc.

Furthermore, each work package (except WP9 management) may organise additional face to face meetings which are to be attended as required and appropriate. The agenda of such meetings can be to address specific matters of relevance in individual work packages, co-operation between work packages etc. Such meetings are decided by the work package leader(s) according to necessity, e.g. at the point close to the completion of an important task/deliverable.

Apart from face to face meetings, there will be telephone conferences at regular intervals among members of each work package. While this can be left to the discretion of the work package leader, it is advised to have a call at monthly intervals, and these calls shall have public minutes. There shall also be regular telephone conferences for the technical board in which the project co-ordinator and the scientific director receive updates on progress and activities from each work package leader. Technical board calls shall occur 4 times a year, between the quarterly face-to-face meetings. Finally, the management board shall have at least one telephone conference a year in addition to face-to-face meetings following the annual review report, plus additional calls as called for by the project co-ordinator.

The project's project coordinator has the powers to call an extraordinary meeting should he feel this is required for the successful continuation of the project or individual parts thereof by giving the partner(s) in question a minimum of 3 weeks notice before any such meeting. In an extreme situation, it may even be necessary for the project coordinator to call an extraordinary meeting of a work package or the entire project. Such meetings will generally be in the form of an extra telephone conference, though a face to face meeting may also be chosen if it is thought to be more effective.

### 2.1.4. Decision process

The decision making process in LinkedTV is organized hierarchically and has as its aim to ensure that issues are addressed at the appropriate level, generally with those "closest" to the issue being tasked with the decision at first, and only in the case of a lack of resolution, shifting the decision process to a higher level. Another aspect of the decision making process is the documentation of decisions which can be of great importance later in the project.

At the lowest point, issues may arise within tasks and should first be resolved between those partners participating in the task, with a final decision being made by the task leader. The issue and decision must be communicated to the work package leader for acceptance. If the work package leader rejects the decision then it must again be addressed by the task participants or may be addressed at the work package level (i.e. involving all work package participants).

At the next point, an issue may be raised at the work package level, e.g. dependencies between tasks, or issues which exist in relation to other work packages. Such issues should be discussed at a work package meeting and the work package leader makes a final decision. One possible decision is to address the matter to the technical board.

The technical board is the next level of decision-making and may have issues referred to it from individual work packages. It is also the right level for resolving issues between work packages. Decisions made at the work package level may be reported to this board, giving other work packages the chance to respond. The technical board can decide to call upon an external expert, e.g. a member of the advisory board, if necessary.

Finally, for issues which may have an impact on the overall project, the management board is the final point for raising and resolving any issue. If necessary, the project coordinator can choose to address an issue at the management level to the EU as final arbitrator.

For documentation of the decision process, technical and management boards will maintain protocols and minutes of meetings, where issues on the agenda will be listed and decisions referenced to the appropriate issue, which shall include the results of any vote as well as subsequent action items for persons with deadlines. Such protocols shall also be maintained by each work package for their records. Out-of-meeting decisions, e.g. at the task level, shall be formally reported by e-mail to the work package leader and added to the next meeting protocol.

## 2.1.5. Conflict and risk management

It is recognized that in the activities of the project, and particularly in the case of hindrances to completion of work, conflicts may arise and it is of great importance to identify, acknowledge and resolve such conflicts as soon as possible.

If a problem is reported by a partner, it is first addressed to the project coordinator. It is also the responsibility of those coordinators to recognize and respond to emerging problems within the project, in order to potentially resolve them before they become more critical. The project coordinator supported by the scientific director will be responsible for identifying issues which may have an impact on the scientific and technological results of the project and taking necessary measures to resolve them. The project coordinator will be responsible for identifying issues which may have an impact on the management of the project and taking necessary measures to resolve them.

When a problem has been identified, the first step is to raise this on the agenda of the next board meeting. If the next board meeting is scheduled to take place at a time which is judged by the coordinator to be too late to begin a successful resolution of the problem, the coordinator may call an extraordinary meeting which can be scheduled to take place with a minimum of three weeks notice. Solutions shall be decided at board meetings using the decision process described above.

Some problems can not be avoided and LinkedTV is also careful to make contingency plans and implement risk management within its management structure. While not all problems can be foreseen, certain common types of problems can be identified, and contingency planning made in advance, as shown in the below table, will be carried out throughout the project's duration.

**Table 8: Risk management**

| Risk | Potential for risk | Impact on project | Contingency plan |
|------|-------------------|-------------------|-----------------|
| Consortium – a partner can no longer provide | Medium – personnel changes can not be avoided but all partners | High – work may no longer be done on time | The LinkedTV consortium is prepared and agreed to |

| | | | |
|---|---|---|---|
| the resources or skills foreseen in the workplan | are very committed to the research themes of the project | | immediately react to this risk by shifting resources internally to partners who can replace the lost resources or skills. All main skill needed for the project are available at more then one partner. |
| Technology – new commercial results are made public which parallel LinkedTV work | Low – while *similar* results may be launched, it is unlikely others will be able to replicate the expertise brought together in LinkedTV to realise truly semantic "meaningful multimedia" services | High – work done may not be able to justify the originally foreseen added value | While we do not expect commercial organisations to be able to replicate LinkedTV work due to the specialised expertise required, technology tracking will take care to pre-empt commercial releases and, if necessary, to re-focus goals of LinkedTV |
| Workplan – the estimated resources prove too little for a task or partner to fulfil the work foreseen | Medium – despite partner experience in these areas of research, it is unavoidable that some work may prove more resourceful than originally thought | High – the LinkedTV workplan is tightly related so the completion of individual tasks is part of the successful completion of the entire project | The project coordinator and management board are responsible to monitor this and some resources may be saved from other tasks where the allocated resources proved too high. Where additional resources are not available, some results may have to be produced outside of the project. |
| Scientific work – a certain task can not achieve its intended goal | Medium – partners have strong experience in the research area of their tasks and have taken care to set realistic goals, however LinkedTV does aim at being innovative and stretching research beyond the state of the art. | Medium – for some tasks, partial results may be able to act as sufficient input for ongoing work. In other tasks, LinkedTV may choose as an alternative to take other solutions which can provide at least part of the required results. | LinkedTV has a consortium of partners with significant research work in the relevant areas and, if necessary, parts of the LinkedTV work could be replaced by prior results (e.g. from previous projects) in order to ensure that the remaining work can continue and produce innovative results |
| Exploitation – LinkedTV results fail to be taken up in the market | Low – while an innovative result does not always guarantee commercial success, LinkedTV aims at achieving a multimedia delivery platform which is clearly a valuable | Medium – while immediate exploitation of the results is the goal of LinkedTV, it may be that the technology is still too immature at the end of the project for commercial uptake, or | The exploitation task runs throughout the LinkedTV project and is in a strong position to measure interest in the broadcasting and WebTV communities through our industrial |

| | goal for broadcasters and WebTV producers. The scenarios ensure that some public broadcasters in Europe will showcase the results. | that commercial broadcasters are still not ready to use such technology (depends heavily on the general uptake of semantic technology in the next 5 years) | partners. It is also able to quickly respond to changes in market outlook, as well as work strongly to influence that market itself with LinkedTV dissemination and standardisation activities. |
|---|---|---|---|

## 2.1.6. Quality Management

Even in the smooth day-to-day running of the project, the work being done must not just be accepted as "done" at any level of quality, but the project management structure will also take care to monitor quality of work being done, identify cases of insufficient quality and react to them early, so that project activity not only stays on time but also retains a satisfactory level of quality.

To achieve this, we implement structures for the quality proofing of work being produced at predefined intervals in the project, allowing in our time plan a period for a quality report, a revision according to this report and a final check before "acceptance" within the project. For reports destined to be read external to the project, e.g. deliverables which are sent to the EU, this quality approach will be enforced, with a first review by an internal reviewer (a member of a partner in the consortium who did not participate in the work) followed by revision and review by the work package leader, and then finally a last request for changes by the scientific director before submitting the deliverable to the EU. Furthermore, for internal reports such as task reports, in the case where subsequent tasks have a dependency on this task, a slightly reduced quality approach is fixed, which shall ensure that task results are satisfactory for the use in subsequent tasks. The work plan is organized in a way that all work done will be made available in either a task report or a deliverable, ensuring that all results, being produced throughout the duration of the project, undergo quality proofing.

## 2.1.7. Website and internal communication

As foreseen in the dissemination task, LinkedTV will set up a website to work as the principle communication point externally to interested academic institutions, commercial organisations and the Web community. In mind of the different target groups, the website will not only offer general information about the project but also targeted content to:

- researchers in the areas covered by the proposal. We expect LinkedTV to be a leading activity in these areas over the next four years. It will actively outreach and network in those research communities

- industry in the broadcasting and Web video sectors. We expect the work of LinkedTV to be of great importance to content owners, network administrators and service providers. Hence we will use the website to generate interest in our work and provide demos and showcases after the appropriate milestones. Through the website, we complement the dissemination tasks and build an industry network interested in the LinkedTV vision.

- Web surfers, who are increasingly using video sites such as YouTube or watching TV broadcasting over the Internet. Through non-technical showcases of the technology and availability of beta releases of the open source LinkedTV platform, we also plan to use the Website to build an active user community which can provide feedback on our

research and form a growing user base (through word of mouth) – an important prerequisite to win the interest of commercial enterprises.

Internally, we will make available a content management system (CMS) where documents produced in the project can be stored and worked on collaboratively between partners. We will use this system for task reports, deliverables (in development), meeting minutes and other internal documents. Documents which shall be made available to the public will, once essential quality levels have been met, be made available for download on the website.

Additionally a software development platform will be used in the implementation of the LinkedTV framework. While later an open source version of the software could be placed on a public site such as Sourceforge, initially we will develop the software privately using a collaborative software development tool such as TraC which integrates Subversion for code control and versioning with a ticket system, wiki and discussion forums.

It is not yet decided if further communication structures will be needed or are desirable. For example, both blogs (for the external presentation) and wikis (potentially for internal co-operation) can be used in LinkedTV but are highly dependent on regular usage by the consortium.

### 2.2. Individual Participants

1. **Fraunhofer IAIS**

Fraunhofer is the largest application oriented research and innovation organization in Europe. It comprises 14000 researchers and engineers in 57 institutes with an overall annual budget of 1.4 Bill €. The Fraunhofer Institute for Intelligent Analysis and Information Systems (IAIS) is the leading Fraunhofer institute in the area of intelligent information systems with special emphasis on business processes, web based multi-media applications, and knowledge mining.

IAIS possesses a broad spectrum of scientific skills and competences needed for the LinkedTV project like web-based multimedia search engines, automated analysis of sound and video recording, speech processing, document analysis and image/video analysis. These skills have been acquired in many years of research and industry activities including EU research projects under the 6th and 7th Framework. In the large scale German Theseus program Fraunhofer IAIS develops semantic search technology and applications in the Core Technology Cluster and for the use case CONTENTUS focusing on intelligent multi media technology applications.

In addition, IAIS is the coordinator of the LIVE EU project and of the FP6 Integrated Project IRRIIS. Furthermore, IAIS is partner in the EU FP6 and FP7 projects Boemie, Vitalas, CitizenMedia, MoveOn, DIESIS and CHORUS.

**Dr. Rüdiger Klein** has a long record of research in areas like intelligent systems and Semantic Web applications. After his Diploma and doctoral thesis at Humboldt University in Berlin he worked at the Academy of Sciences in Berlin in the Knowledge Based Systems Lab. Rüdiger Klein joined Daimler Research in Berlin 1991 as a senior researcher and project leader with a focus on intelligent system applications in industrial environments. He was responsible for knowledge management systems with Mercedes and Airbus. He was a member of the EU Project MOKA and of the industrial advisory boards of EU Projects IBROW and OntoWeb. Two years ago Rüdiger Klein left Daimler Research and joined Fraunhofer IAIS. He is the project coordinator of the EU Integrated Project IRRIIS.

**Dr. Joachim Köhler** received his diploma and Dr.-Ing. degree in Communication Engineering from the RWTH Aachen and Munich University of Technology in 1992 and 2000, respectively. In 1993 he joint the Realization Group of ICSI in Berkeley where he investigated robust speech processing algorithms. From 1994 until 1999 he worked in the speech group of the research and development centre of the SIEMENS AG in Munich. The topic of his PhD thesis is multilingual speech recognition and acoustic phone modelling. Since June 1999 he is with Fraunhofer IAIS in Sankt Augustin and head of the department NetMedia. The research focus of NetMedia lies in the area of multimedia indexing and search methods and applications. His current research interests include pattern recognition and speech recognition, spoken document and multimedia retrieval and multimedia information systems. He is currently involved in the European research CHORUS project which creates a roadmap for audio-visual search technologies and is technical co-ordinator of the European IP-project LIVE.

**Dr. Jobst Löffler** has been working as research scientist at the Fraunhofer IAIS since 1997 in the areas of distributed media information systems, automatic media analysis and software architectures for cooperative environments and digital libraries. He was responsible for the EU project SHARE as technical coordinator and project leader of the iFinder project which aims at developing a media analysis and retrieval system as a product for the media industry.
In 2002 he received his PhD in computer science from the Technical University of Braunschweig, Germany with a dissertation about adaptive visualization of 3D documents for cooperation within open information spaces.

2. **CERTH-ITI**

**Centre for Research and Technology Hellas – Informatics and Telematics Institute (CERTH-ITI)**

The Informatics and Telematics Institute (http://www.iti.gr) of the Centre for Research and Technology Hellas (http://www.certh.gr) was founded in 1998 as a non-profit organisation under the auspices of the General Secretarial of Research and Technology of Greece (GSRT), with its head office located in Thessaloniki, Greece. Since 2000 it has been a founding member of the Centre of Research and Technology Hellas (CERTH) also supervised by the GSRT. CERTH-ITI's related areas of R&D activities include: semantic multimedia analysis, multimedia indexing and retrieval, large-scale and social media analysis, knowledge structures, languages and reasoning for content analysis, personalization and knowledge discovery for Semantic Web applications, intelligent human computer interaction. CERTH-ITI has participated in more than 50 EC IST and 85 National projects and subcontracts. Over the last eight years, the ITI research team has authored over 200 publications in scientific journals, 65 books and book chapters and over 500 presentations to international conferences.

CERTH-ITI currently coordinates the FP7 ICT IP WeKnowIt "Emerging, Collective Intelligence for personal, organisational and social use" and has leading roles in the FP7 ICT IP GLOCAL "Event-based Retrieval of Networked Media", FP7 ICT STREP JUMAS "Judicial Management by Digital Libraries Semantics", and FP7 ICT STREP PESCaDO "Personalized Environmental Service Configuration and Delivery Orchestration". It also participates in the European Defense Agency funded project MEDUSA "Multi Sensor Data Fusion Grid for Urban Situational Awareness", and in the audiovisual search engines Coordinated Actions "CHORUS" & "CHORUS+". In the past, CERTH-ITI participated and had leading roles in several IST-FP6 IPs (aceMedia, MESH, X-Media), IST-FP6 STREPs (BOEMIE, VIDI-Video) and NoEs (Knowledge Web, K-Space), and also lead the Greek national R&D Training Network MULTI-MINE. CERTH-ITI also participates in standardization activities, such as activities carried out within the MPEG, JPEG and W3C standards groups, and is very active in the organization of related scientific conferences,

including the 2nd ACM Workshop on Events in Multimedia (EiMM 2010), the 2nd ACM Workshop on Social Media (WSM 2010), ACM CIVR 2009, and WIAMIS 2007.

**Dr. Ioannis Kompatsiaris** is a Senior Researcher (Researcher B') with the Informatics and Telematics Institute of CERTH. His research interests include semantic multimedia analysis, indexing and retrieval, Web 2.0 content analysis, knowledge structures, reasoning and personalization for multimedia applications. He is the co-author of 39 papers in refereed journals, 20 book chapters and more than 150 papers in international conferences.

**Dr. Vasileios Mezaris** is a Senior Researcher (Researcher C') with the Informatics and Telematics Institute of CERTH. His research interests include image and video analysis, content-based and semantic image and video retrieval, knowledge-assisted multimedia analysis, knowledge extraction from multimedia, medical image analysis, multimedia standards,. He is the co-author of 17 papers in refereed international journals, 6 book chapters, and more than 65 papers in international conferences.

## 3.  University of Economics Prague

The group at UEP is recognised for its research and educational activities in knowledge discovery from databases, web/text/multimedia mining, web engineering and knowledge-based systems. It recently participated as funded partner in seven EU projects: in the multimedia area (6FP NoE K-Space), in the knowledge discovery from databases area (5FP projects Sol-Eu-Net and MiningMart), in the medical informatics area (4FP project MGT and DG SANCO project MedIEQ), in the e-learning area (6FP IP KP-Lab), and in the digital libraries area (eContent M-CAST). The group was involved in multiple EU network projects such as KDnet, Knowledge Web, Ontoweb or EUNITE, and its members participate in several W3C working groups. The group also host/ed top-class international conferences such as ECML (1997), PKDD (1999), EKAW (2006) and ISMIS (2009). The key expertise wrt. LinkedTV is related to *web mining and information extraction*, as well as to the issues of *multimedia-text analysis complementarity* and *multimedia metadata systems*, which will be exploited in WP1 and WP2 of LinkedTV. Web mining and information extraction was thoroughly studied in the EU MedIEQ project and in the national (CSF-funded) project Rainbow. In the EU K-Space project, in turn, UEP has been the leader of the task devoted to mining complementary resources to multimedia, and contributed to the development of the COMM – core ontology for multimedia. From KP-Lab UEP will bring the experience with text-based construction of shared domain models, and from M-CAST that with online question answering.

**Dr. Vojtěch Svátek** obtained the PhD in Informatics from the UEP in 1998 and became Associate Professor in 2007. His main research domains are data/text/web/multimedia mining and ontological engineering. Local contact person in EU-funded projects K-Space and MedIEQ, co-ordinator of two grants of the Czech Science Foundation. Program Co-Chair of EKAW 2006 conference, PC member of about ten other relevant conferences (ECML/PKDD, ESWC, ASWC, SAMT, BIS, CBMI etc.). Member of the W3C OWL working group.

**Prof. Petr Berka** obtained the PhD in Bionics from the Czech Technical University in 1991, became Associate Professor in 1995 and Full Professor in 2005. His research and educational activities concentrate on machine learning, data mining and rule-based systems, with applications in medicine, finance and information retrieval. Local contact person in the EU-funded project MiningMart. Program (Co-)Chair of the ISMIS 2009 conference. Member of ECCAI and SIGKDD. Most recently he led experiments on applications of fuzzy rule reasoning for semantic region merging in multimedia analysis (within the K-Space project).

## 4.  Centrum Wiskunde & Informatica (CWI)

The Interactive Information Access group at the Centre for Mathematics and Computer science (CWI) carries out research on improving models and tools for presenting multimedia information to end-users on a variety of platforms. CWI is the research institute for mathematics and computer science research in the Netherlands. CWI's mission is twofold: to perform frontier research in mathematics and computer science, and to transfer new knowledge in these fields to society in general and trade and industry in particular. CWI has always been very successful in securing considerable participation in European research programs and has extensive experience in managing these international collaborative research efforts. CWI is also strongly embedded in Dutch university research: about twenty of its senior researchers hold part-time positions as university professors and several projects are carried out in cooperation with university research groups. In addition, CWI has strong links to the World Wide Web consortium, and houses the Benelux office. CWI has a staff of 210 fte (full time equivalent), 160 of whom are scientific staff. CWI operates on an annual budget of EURO 13M.

**Prof. dr. Lynda Hardman** is the head of the Interactive Information Access group and part-time professor at the University of Amsterdam. She obtained her PhD from the University of Amsterdam in 1998, having graduated in Mathematics and Physics from Glasgow University in 1982. During her time in the software industry she was the development manager for Guide, the first hypertext authoring system for personal computers (1986). She participated in the EU FP6 K-Space Network of Excellence and several other EU research projects in the past. She co-edited a special issue of the Multimedia Systems Journal on the canonical processes of media production and a special issue for IEEE Intelligent Systems on AI and Cultural Heritage. She was co-programme chair for the international conference on semantic and digital media technologies (SAMT) in 2008. She is a member of the editorial board for the Journal of Web Semantics and the New Review of Hypermedia and Multimedia. Her favourite chocolates are from Puccini, Amsterdam.

**Dr. Jacco van Ossenbruggen** is a senior researcher at CWI and a part time assistant professor at the Vrije Universiteit of Amsterdam. He obtained his PhD from the Vrije Universiteit Amsterdam in 2001. He was a member of the W3C working group that developed the SMIL recommendation. He was deputy project manager of the Dutch MultimediaN E-culture Project, which won the first prize at the Semantic Web Challenge at ISWC'06. He participated in the EU FP6 K-Space Network of Excellence and several other EU research projects in the past. He is currently working on the EU FP7 PrestoPrime and EuropeanaConnect projects and on the Europeana Thoughtlab demo. His current research interests include multimedia on the Semantic Web and the exploration of heterogeneous media repositories.

**Prof. dr. ir. Arjen P. de Vries** is a senior researcher at CWI and a part-time full professor in the area of multimedia data spaces at the Technical University of Delft. He received his PhD in Computer Science from the University of Twente in 1999, on the integration of content management in database systems. He is especially interested in the design of database systems that support search in multimedia digital libraries. He has worked on a variety of research topics, including (multimedia) information retrieval, database architecture, query processing, retrieval system evaluation, and ambient intelligence. He coordinates the TREC and INEX Entity Ranking tracks, which is of particular relevance to the search activities in this project. In 2004, De Vries and his then PhD student Westerveld received the best paper award in the international conference on image and video retrieval (CIVR), and in 2007, De Vries and his PhD student Cornacchia received the best student paper award in the European conference on Information Retrieval (ECIR). He is currently participating in the EU projects Vitalas (FP6) and PuppyIR (FP7).

5. **Eurecom**

Eurecom, Sophia Antipolis, France, is a graduate education and research center, funded by two schools: Telecom ParisTech (France) and EPFL (Lausanne, Switzerland), with several academic and industrial members. Our research activity is organized in three themes: mobile, corporate and multimedia communications. We have a very active collaboration program, and participate in many projects at the national (Argos, RPM2) and European (STATION, GM4iTV, PorTiVity) level. Research at Eurecom includes in particular topics such as signal processing, information theory, speech processing, watermarking, biometry, multimedia analysis, information filtering. Our group is a regular participant in the TRECVID evaluation campaigns. We have participated in several European projects, including the K-Space Network of Excellence. We recently organized the 15th Multimedia Modeling international conference (MMM 2009) in January 2009.

**Dr. Raphaël Troncy** is currently Assistant Professor in the multimedia information processing group of Eurecom (France). He obtained with honors his Master's thesis in Computer Science at the University Joseph Fourier of Grenoble (France), after one year spent in the University of Montreal (Canada). He benefited from a PhD fellowship at the National Audio-Visual Institute (INA) of Paris where he received with honors his PhD from the University of Grenoble (INRIA/INA) in 2004. He selected as an ERCIM Post-Doctorate Research Associate 2004-2006 where he visited the National Research Council (CNR) in Pisa (Italy) and the National Research Institute for Mathematics and Computer Science (CWI) in Amsterdam (The Netherlands). He was a senior researcher for CWI from 2006 till 2009. Raphaël Troncy is co-chair of the W3C Incubator Group on Multimedia Semantics and the W3C Media Fragments Working Group, contributes to the W3C Media Annotations Working Group and actively participates in the EU K-Space Network of Excellence. He is an expert in audio-visual metadata and in combining existing metadata standards (such as MPEG-7) with current Semantic Web technologies. He works closely with the IPTC standardisation body on the relationship between the NewsML language family and Semantic Web technologies.

**Dr. Benoit Huet** received his BSc degree in computer science and engineering from the Ecole Superieure de Technologie Electrique (Groupe ESIEE, France) in 1992. In 1993, he was awarded the MSc degree in Artificial Intelligence from the University of Westminster (UK) with distinction, where he then spent two years working as a research and teaching assistant. He received his DPhil degree in Computer Science from the University of York (UK) for his research on the topic of object recognition from large databases. He is currently Assistant Professor in the multimedia information processing group of Eurecom (France). His research interests include computer vision, content-based retrieval, multimedia data mining and indexing (still and/or moving images) and pattern recognition. He has published over 80 papers in journals, edited books and refereed conferences. He is a member of IEEE, ACM and ISIF. He has served in many international conference organization and technical program committee. He is regularly invited to serves as reviewer for prestigious scientific journals as well as expert for project proposal at national, European and International level. He is the conference chair of the International Conference on Multimedia Modeling (MMM'2009) which took place in Sophia-Antipolis (France) in January 2009.

**Prof. Bernard Merialdo** was admitted in the Ecole Normale Supérieure (Maths section) in 1975. He received a Ph.D. in Computer Science from Paris 6 University in 1979 and an "Habilitation à Diriger des Recherches" from Paris 7 University in 1992. He first taught at the Faculty of Sciences in Rabat (Morocco). In 1981, he joined the IBM France Scientific Center in Paris, where he led several research projects on natural language processing and speech recognition using probabilistic models. From 1988 to 1990, he was a visiting scientist in the IBM T.J Watson Research Center in Yorktown Heights, N.Y. (USA). In 1992, he became a professor in the Multimedia Communications Department of Eurecom. His current research topics are multimedia indexing and information filtering applications. He is a member of IEEE, ACM, he was associate editor for the IEEE Transaction on Multimedia, and general chair for the ACM Multimedia 2002 conference. He participates in several conference program

committees and expert boards. He is currently Head of the Multimedia Communications Departement at Eurecom.

### 6.   University of St Gallen

**University of St. Gallen** is the university for Business Administration, Economics, Law and Social Sciences, founded in 1898 as a "business academy", is one of the oldest universities of its kind in the world. The University of St. Gallen today is one of Europe's leading business schools and was among the first to be accredited by international EQUIS and AACSB standards. In this setting the mcm institute – founded in 1998 with support of the Bertelsmann Foundation and the Heinz-Nixdorf Foundation – has established itself as a leading international centre in the field of media and communication for research and consulting. The institute is exploring business models, acceptance and the management of new media as well as the application of digital media in corporate communication. With its engagement in leading cooperative research projects, the University of St. Gallen has broad access to experts and knowledge in the field of mobile service creation. Furthermore, the mcm institute-Forschungsbeirat (Advisory Board) equipped with C-level business- and thought-leaders from leading ICT and media companies (e.g. Swisscom, SAP AG), guarantees vital exchange of ideas for applied research. In addition, collaborations with worldwide leading business schools (e.g. MIT, Columbia University of St. Gallen) fertilise the research on the forefront of management science.

**Prof. Dr. Katarina Stanoevska** is vice-director of the =mcminstitute and associated professor at the University of St. Gallen. Her research work is focused on business models and strategies in telecommunication and media industries. She is in particular interested in topics related to market trends, innovative advertising and communication models based on connected media as well as development and evaluation of business models. She has published numerous papers in these areas and has participated as work package leader for business aspects and user requirements and as exploitation manager in four European projects related to mobile services. From 1992 till 1997 she worked as a Research Assistant and Doctoral Student at the Institute for Information Management, University of St. Gallen, where she also got her Ph.D.. In1997 she was a Project Manager at the Institute for media and communication management. Since September 2004 she becomes a member of the executive board of the mcm institute and since the beginning of 2005 assistant professor and vice-director of the Institute for Media and Communications Management. Since January 2009 she is elected associate professor at the University of St. Gallen.

**Thomas Wozniak** works as Research Assistant and is doctoral candidate at the Institute for Media and Communications Management of University of St. Gallen (Switzerland). He has been involved in EU research projects in the fields context-aware mobile services, thereby focusing on business aspects. He works on market analyses, business models, and provides business consultancy to project partners.
Thomas Wozniak studied at University of Leipzig (Germany), Berlin University of the Arts (Germany), and Edith Cowan University Perth (Australia). He holds a diploma degree in Business Administration of the University of Leipzig and a diploma degree in Electronic Business of the Berlin University of the Arts.

### 7.   Semantic Technologies Institute International

STI International was founded in April 2007 and is established as a research association within Austria. As a natural outcome of the European Semantic Systems Initiative ESSI, the EU FP6 Network of Excellence Knowledge Web, and DERI International, STI International is a mature association of interested scientific, industrial and governmental parties sharing common R&D objectives: to establish semantics and semantic technologies as an integral part of modern

computer engineering. Currently we can count 25 public and corporate members, whilst several further member applications are being processed. To support its members and promote semantics, STI International coordinates and actively contributes to major research and education activities in Europe and promotes greater awareness and faster take-up of semantic technology in full synergy with these activities. The association is structured accordingly in three areas (*research*, *technology*, *realization*) which further comprise a series of services to the members as well as to other external parties (*education*, *road maps*, *commercialization*, *standardization and reference architectures*, *test beds and challenges*). Designated area leaders and service managers are responsible for the successful operation of STI and its services.

**Alexander Wahler** is CEO of STI International. He is also co-founder of Hanival Internet Services GmbH and has a degree in electrical engineering from the Vienna University of Technology. He has many years of experience in the management of European research projects and is currently leading the work of Hanival in the IST-FP6 project SUPER. He was an advisor to the Austrian Chamber of Commerce on the FP7 Framework programme.

**Dr. Lyndon J B Nixon** has joined STI International as senior postdoctoral researcher in November 2008. Previously he was a researcher at the FU Berlin, where he acted as Industry Area Co-Manager of the EU NoE KnowledgeWeb and double Workpackage Leader in the EU STREP TripCom. In KnowledgeWeb, Dr. Nixon organized and led activities promoting the transfer of semantic technology to industry. He received his PhD in January 2007 with the topic 'Semantic Web enabled Multimedia Presentation system'. His research focus is Web-based TV/video and the semantically guided integration of Web-based content, and he has organized a number of workshops around related themes.

**Katharina Siorpaes** is a researcher at the Semantic Technology Institute and at the start-up company playence. She holds a master of Computer Science from the University of Innsbruck and is currently finishing her PhD (Q2 2009). Her main research interests are community-grounded evolution of ontologies, annotation of multimedia, incentives for semantic content authoring, and games for semantic content creation. In her PhD research, Katharina has created OntoGame, games for building ontologies and annotating multimedia content. Katharina is co-organizer of the workshop on incentives for the Semantic Web (INSEMTIVE'2008) at ISWC 2008 at of the workshop on incentives for the Web (WEBCENTIVE'09) at WWW'09.

## 8. Condat AG

The Condat AG is a medium sized company located in the centre of Berlin developing and integrating innovative solutions for leading European companies. Condat is one of the main german providers for planning, distribution and media asset management including EPG for the major public and private TV-broadcasters (e.g. RBB, ARD, MDR, WDR, NDR3, n-tv, arte, Deutsche Welle). Condat has actively participated (co-ordinating or members of consortiums) in 7 EU-Projects from the 3rd to the 6th Framework Programme in different domains concerning TV / New Media, Semantic Web and Mobile Applications.
Condat provides Web-TV solutions introduced by television broadcasters to offer their video materials on-demand via the internet. The attractive presentation of the videos requires to include a comfortable navigation, advanced search, meta data and user profile evaluation. We apply semantic search engines to retrieve and analyse large, heterogeneous data sources distributed throughout the network for the semantic connection of objects and the ensurance of integrity and actuality of links. Our business unit mobile applications develops personalised user interfaces for different mobile devices using descriptions concerning screen format, in-/output facilities and operational system. Several mobile TV solutions have been developed, which show video transmitted via WLAN, DVB-H or DVB on mobile devices.

Condat covers the entire Plan – Build – Run cycle applying modern methods such as V-Model, OOP. The development of server and client side applications uses J2EE, Java, Web- and Open Source – technology. Project and quality management is certified according to DIN EN ISO 9001:2000. In the LinkedTV project Condat contributes in the areas architecture specification and implementation, network search, semantic classification of static and dynamic content, profiling of programs for different user profiles. Due to our close relation to customers in the TV sector, we participate in the field trials to validate the project approach and pilot system.

**Dr. Christof Peltason** is COO at Condat AG. He is in charge of large industrial customer projects for the TV and media sector. His main interests and activities in the past have been the development of TV/Media solutions, corporate portals, content management and semantic web based platforms. He has studied mathematics and computer science at the University of Bonn and holds a PhD from Technical University Berlin with the subject "Knowledge Representation for Systemdesign: the Management of Classifications and Taxonomies" (1989).

**Rolf Fricke** graduated in Diplom-Informatik (computer science) in 1985 at the Technical University of Berlin. His focus is the design and development of semantic web based Internet- and mobile applications on the basis of application servers, search engines, web content management (CMS) and Web-TV portals. He was involved in several industry and research projects and is familiar with Object-oriented methods for analysis and design (OMT, UML), database design and development.

**Kerstin Mathaj** graduated in Diplom-Informatik (computer science) at the Technical University of Berlin in 2004. Her focus is the design and development of model-driven Video-, TV- and mobile applications on the basis of application servers, web content management (CMS) and Semantic Web technology. She is familiar with Object-oriented methods for analysis and design (OMT, UML), SOA (Service Oriented Architectures), database design (Oracle, MySQL) and network search (Lucine).

9. **Noterik**

Noterik, established in 1996 in Amsterdam, the Netherlands, is a privately held company (SME). The founders, (MSc) mr. J. Gural and mr. N. Hershler, started the company as Noterik & Doonder V.O.F. The legal identity of the company was changed into B.V. (corresponding to Ltd) in 2001.
Noterik is an R&D driven full service multimedia company providing efficient solutions in the field of digital media, employing highly skilled software engineers and (interaction) designers which develop media applications based on open standards. Noterik operates with a multilingual staff of approximately 30 people.
The solutions Noterik develops in the field of indexing, management, retrieval and delivery of digital media are related to the StreamEdit framework, a Media Content Management application (MAM), also known as a Digital Asset Management system (DAM) developed by Noterik. Clients and partners include European broadcasters, the European Commission, National Research and Educational Networks (NERN's), large public service organisations and well known national and international research organisations. Noterik has received a number of awards including the Dutch Telework (2002) and the Dutch Europrix (2004) award.

**Rutger Rozendal** is the Chief Financial Officer of Noterik. He started working at Noterik in 2003. During the first years he was responsible for the coordination and implementation of webcasting services to different customers. In 2006 the service expanded to include educational webcast functionality so that the technique could be used by learning institutions to provide online interactive lessons. Rutger oversaw implementation of this technology in both the medical and financial sectors. Since 2009 Rutger is responsible for the overall

marketing and sales strategy of Noterik. In close collaboration with Daniel Ockeloen, Head of the Technical Development, he provides the roadmap for the WebTV development for the coming years.

**Daniel Ockeloen** started as software designer. In 1993 he was involved in writing software for the first ISP in the Netherlands, During this time he was asked by one of the Dutch National Broadcasters (VPRO) to setup their digital department, From 1994 to 2000 he was responsible for over 30 projects ranging from live media events to designing applications. One of the products was MMBase a CMS designed for national broadcasters, he prepared MMBase for open-source release in 2000, he left VPRO to co-found the multi-media company Submarine, which produces cross-media products where he took the role of CTO. Mr. Ockeloen was part of the MMBase Management Committee for the first few years of the open-source phase. In 2004 he left Submarine to work on MMBase fulltime mostly hired by Dutch broadcasters and education related companies. In 2006 he was asked to join Noterik as Head of the Technical Development.

**Rutger Bückmann** has a degree in cognitive artificial intelligence (MSc.). He works as a project leader/software architect for Noterik. He is a specialist in software engineering and has a lot of experience with the development of Java applications and webservices that deal with video and 2D animation. He is involved in the development of E-learning applications for Noterik that have been in use by teachers and students from various universities in the Netherlands. He is leading a workpackage on Content Composition, Syndication and Usage in the FP6 IST MESH ([www.mesh-ip.eu](www.mesh-ip.eu)). Furthermore is extensive knowledge of audio processing and algorithmic music composition.

## 10. The Netherlands Institute for Sound and Vision (NISV)

The Netherlands Institute for Sound and Vision (NISV) maintains and provides access to 70 per cent of the Dutch audio-visual heritage, comprising approximately 700,000 hours of television, radio, music and film, making NISV one of the largest audiovisual archives in Europe. NISV combines the highest professional standards concerning the release and storage of material, with easy access for its users, by using state of the art systems for asset management and storage.
NISV is the business archive of the national broadcasting corporations as well as a cultural heritage institute. NISV has brought thousands of hours of archive footage on-line for educational use and also operates a facility for the general public, the Media Experience, which is visited by 200,000 people annually. In 2007, the seven-year Images for the Future programme was launched. The project, funded with a budget of €154 million from the FES Fund, will realise the digitisation of the Netherlands' audiovisual memory in a process whereby NISV will conserve and digitise 17,500 hours of film, 124,000 hours of audio, 137.000 hours of video and 1.2 million photos. This material will be made available for (broadcast) professionals, education and the general public. In the scope of this prioject, Sound and Vision is experimenting with crowdsourcing.
NISV is an experienced partner in National and European research projects and active in the following international organisations FIAT/IFTA, EBU and UNESCO. Current research projects include LiWA, PrestoSpace,MultiMatch, VidiVIDEO, P2PFUSION, Video Active, COMMUNIA, CATCH programme, and MultiMediaN.

**Johan Oomen** is head of the NISV He is mainly working on externally funded research projects. Since June 2007, he joined the Images for the Future project team. He is part of the research groups of national and international research projects. He holds a BA in Information Science and an MA in Media Studies. He is member of the Webstroom expert group funded by the SURF Foundation, on the use of streaming media in higher education and general secretary of the DIVERSE network. Johan Oomen has also worked for the British Universities Film and

Video Council and the Holland Media Group. He has given papers at leading conferences and published several articles in journals, including Ariadne, Innovate and Informatie Professional. His book Internet en het Nieuwe Leren: de toepassing van streaming media was published recently.

**Roeland Ordelman** is project manager R&D at the Netherlands Institute for Sound and Vision and researcher Speech & Language Technology and Multimedia Retrieval at the University of Twente. He is co-founder of XMI (Cross Media Interaction), a company that provides services for automatic, speech-based annotation of audio data. He received his PhD on "Dutch Speech Recognition in Multimedia Information Retrieval" at the University of Twente in 2004. He has been working on a number of national and international projects in the area of multimedia retrieval and speech and language technology, such as more recently LiWA, MESH, MediaCampaign, MultimediaN and the CATCH project CHoral. His work focuses on deploying multimedia retrieval technology for accessing information in audiovisual data and enhancing the exploitability of the information that is available within and along with audiovisual data collections. He is specifically interested in robust speech recognition, time-synchronization of collateral textual data such as minutes, subtitles and transcripts with audio and the representation of information that is extracted from the audio stream. The latter also includes exploiting user communities for correction/enhancement of automatically generated audio labels.

**Hans Westerhof** is programme director at NISV. He is responsible for Images for the Future, a large digitisation project. Hans Westerhof is also manager of the department of collections, which looks after the total of 700.000 hours of analog and digital collections of the archive. Before his work at the archive he worked at Knowledgeland, an Amsterdam-based think tank, at Andersson Elffers Felix, a management consultancy firm and for the Amsterdam Municipality, as political advisor to the alderman of urban planning and housing.

## 11. University of Mons

The POLETI (Research Center in Information Technologies) group of the Faculty of Engineering at the University of Mons (UMONS), Mons, Belgium, is active in the following areas and their applications: voice quality processing and voice modeling, text-to-speech systems, statistical pattern classification, noise robust speech recognition algorithms, audio processing, automatic character recognition, biomedical data processing, sensor an data fusion, image processing, DSP implementation and software optimization. Activities in the area of multimedia content analysis and human behavior analysis are in particular pursued within the NUMEDIART (http://www.numediart.org) research project, a long term national research program centered in Digital Media Arts. In this project, the POLETI group collaborates with other Belgian partners with developments related to hypermedia navigation in the context of artistic installations or performances, as well as media production needs. Our group also benefits from international collaborations with first class institutions and involvement in several European projects (e.g., HIMARNNET, SPRACH, RESPITE, THISL, SIMILAR Noe, CALLAS IP, COST 232, 249, 250, 278, 2102 and 2103) and national projects. The POLETI group also has R&D contracts with the industry (e.g., ACEC, PHILIPS, SAIT), owns several international patents, and has created ACAPELA S.A, a spin-off company licensing speech technologies (30 employees). The POLETI has also co-created a spin-off R&D Lab, Multitel ASBL, which now has 60 people on its payroll.

**The numediart Institute for New Media Art Technology** was founded in 2010 by the University of Mons. Building upon MONS 2015 (Mons will be the EU Capital of Culture in 2015), the Institute organizes internationally-renowned scientific training and research activities in the area of new media art technology. The topics covered by the Institute are : audio, image, video, gesture, and bio-signal processing, for applications in which man-machine inter-

action aims at creating emotions. Through UMONS, the media arts scenario (WP7c) will be executed in collaboration with the numediart Institute.

**UMONS Key personnel**

**Prof. Bernard Gosselin** is professor of Image and Information Processing at the Faculty of Engineering of the University of Mons. He is also a Research Consultant at Multitel Research Center, where his responsibilities cover image processing and video activities. Since 2001, he has been heading the Image Processing Research Group at the Faculty of Engineering, Mons. He is a member of program committees of several international journals and conferences on image processing, artificial intelligence, and pattern recognition. He is author or co-author of more than 50 international publications in journals and conference proceedings. He has been an invited speaker in several conferences, and visiting professor at the University of Orléans, France. His expertise in image analysis and information processing covers, through several past or ongoing research projects, the fields of pattern recognition, defect detection and classification, traffic analysis and video surveillance, and video or secure signal processing. He has supervised several PhD theses in these fields.

**Dr. Matei Mancas** holds a PhD in applied sciences from the FPMs (now Faculty of Engineering of the University of Mons, Belgium) since 2007. His research deals with signal understanding and computational attention. This mechanism is used by humans in order to select "important" features in an image or a sound and it is the beginning of a higher level signal understanding. He is currently working on the use of computational attention for expressive gestures and emotion description and in social interaction analysis. Dr. Mancas wrote several national and European projects for several years and he has been a member of the SIMILAR network of excellence. He is also a member of Eucognition network and a Belgian expert in a COST action dealing with real-life complex video analysis.

**Dr. Stéphane Dupont** received the PhD degree in Electrical Engineering at FPMs (now Faculty of Engineering of the University of Mons, Belgium) in 2000. He has been post-doctoral associate at the ICSI (California) in 2001-2002. There, he participated to the ETSI standardization activity on robust distributed speech recognition over wireless networks (Aurora). In 2002, he joined Multitel (Belgium), a research and innovation center, to be in charge of speech recognition research. There, he organized the writing and coordinated several projects, including the EU FP6 DIVINES project (multimodal interfaces challenge). He joined UMONS in 2008. Dr. Dupont interests are in multimodal and speech interaction technologies, multimedia content management, computer music, pattern recognition, artificial neural networks and signal processing. He has authored/co-authored over 40 papers on these topics and holds 3 international patents.

## 12. RBB

RBB is the regional public broadcaster for the federal states of Berlin and Brandenburg. It produces and broadcasts one television channel and six radio stations in addition to providing a range of interactive multimedia services including websites, community video platforms, mobile and teletext services. As part of the ARD (Association of Public Service Broadcasting Corporations in Germany), it contributes to several nationwide TV and radio channels and multimedia services. RBB hosts and manages the ARD Play-Out-Center which is responsible for the playout of the ARD Digital programme bouquet via DVB-S/T and the production and editing of SI data for all ARD TV and radio programmes. RBB employs approx. 1,500 staff in addition to a large pool of free-lancers.

RBB has longstanding experience in EU co-funded research projects, gaining extensive knowledge in the creation of innovative formats and services, in content production and distribution and also in usability testing of services. Recent activities concentrated on collaborative content production, personalized media and accessibility services. RBB follows the most recent developments in interactive content authoring technologies, such as Adobe Air, JavaFX, MPEG-LASeR, and BD-Java (BluRay) and is capable of prototyping applications and content based on these technologies for different end devices. Annotation metadata formats such as TV-Anytime, MPEG-21 or APML are also within its research focus.

**Sven Glaser** was awarded his diploma in International Media and Computing from University of Applied Sciences in Berlin (FHTW) in 2005. Since 2006 Sven Glaser has been employed by RBB as project engineer for its Innovation Projects. His special focus is on the DVB framework and its combination with network technologies as well as user-centred interaction paradigms and barrier-free access technologies. He was RBB's technical manager in the FP6 project PORTIVITY and is now working on the PSP project DTV4All and FP7 project VITAL++.

**Nicolas de Abreu Pereira** studied English and History of Art at Freie Universität Berlin. He has worked on several EC projects with a focus on user interactivity, including MHP-KDB – The MHP Knowledge Base, ENTHRONE 2, OnRadio and Audiotain (as project coordinator) for RBB and its predecessor. While working at Berlin's Technical University he concentrated on user requirement analysis in the fields of eGovernment and eParticipation, technologies for the Ageing Society, technology-enhanced and innovative learning concepts. He is currently RBB's project manager in the FP7 project VITAL++, focusing on IPTV over IMS-based P2P networks.

### 2.3. Consortium as a whole

The success of LinkedTV will be guaranteed through the combined expertise and knowledge of its consortium members. As an Integrated Project, we focus on a consortium made up of a significant number of academic research partners, including one economic unit, a couple of SMEs which stand to benefit from LinkedTV market opportunities and several large traditional media players who will benefit from the innovation of LinkedTV technology to strengthen their respective market position (broadcasters, IPTV system provider, telecommunications).

Firstly, as project co-ordinator, a partner is needed who can provide proven project administration skills and experience, so that day-to-day project operation runs smoothly, as well as a clear vision for the aims and outcomes of LinkedTV, so that the research and innovation proceeds according to a clear, directed path to a common, well defined goals. Only under these conditions it is possible to trust in the successful running and completion of LinkedTV. Here, **Fraunhofer IAIS (DE)** demonstrates the required project administration skills and experience, being project coordinator in several past and present EU projects.

To provide the research expertise in content analysis for relations extraction, annotation and adaptation, we have two partners who have clear leadership in this area, including the successful work in these areas in past EU projects. Results from these projects will be acquired and extended for LinkedTV. **CERTH (GR)** is a leading research institute in the area of semantic multimedia, using hybrid low level and high level content analysis approaches to bridge the so-called „semantic gap" and enable conceptual annotation of media streams. The **University of Economics, Prague (CZ)** is a pioneer in the use of data mining to improve media annotation processes, enabling the association of related media through shared concept identification in accompanying data sources (e.g. subtitles on a TV program, surrounding text on a Web page). **Eurecom (FR)** will complement both partner's activities with their

multimedia research unit, and particularly expertise in appropriate multimedia description schemes for referencing media fragments and tying fragments to concepts. Through their combined research expertise, these partners will also create an infrastructure for the „Connected Media" services on the Internet.

This expertise is complemented by a leading institute in multimedia retrieval and presentation research **CWI (NL)**. CWI is a leader in multimedia metadata research, and will support LinkedTV in the use of appropriate, simple metadata schemes for annotating and linking media content. As the developer of the automated presentation system Cuypers and e-culture search portal MultimediaN, its experience in multimedia presentation and interaction will be used in designing the LinkedTV user interface. Research work on multimedia annotation and presentation generation will be complemented by **STI International (AT)** through the senior researcher Dr Lyndon Nixon, whose PhD focused on this topic. Furthermore, as he has instigated the proposal and given the vision for LinkedTV, he has the clear vision and goal for the project which is necessary in a scientific coordinator. Further research support will be provided through the startup Playence as a third party, whose activities focus on user interfaces and capturing user feedback. STI also brings its experience as organizer of leading events in the academic and industrial communities and its worldwide network of experts in semantic technologies, to make it the ideal partner for dissemination activities.

Given the potential of LinkedTV to generate new business models for traditional media companies as well as the need during the project to keep up to date with the rapidly moving IPTV market so that LinkedTV can be best positioned to innovate in this market and introduce a powerful paradigm shift with the potential to disrupt, we draw on the expertise of **University of St Gallen (CH)** to carry out market surveys, develop business models and support our industry partners in the project in best profiting from LinkedTV.

For development of the back-end infrastructure, including metadata repositories and service endpoints, and to support the broadcasters in integrating LinkedTV technologies into their content processes, the SME **CONDAT (DE)** Condat has been chosen as they bring the required expertise and skills to the consortium. Condat has longterm experience regarding the development and customizing of search engines to retrieve and analyse large, heterogeneous data sources. This knowledge will be used in the project to weave video annotations with other media and configure crawlers to maintain the integrity and actuality of all links.

For development of the LinkedTV player, as well as support for video preparation and delivery at the broadcaster side, the SME **Noterik (NL)** is "on board". Noterik bring their ClickVideo technology to the project. Furthermore, they will provide the IPTV client development skills including realisation of interactive services through widgets.

It is of vital importance in LinkedTV to have a tight co-operation with major players in the areas of broadcasting and telecommunication. The results of the project have the potential to provide new services and business models for broadcasters (offering their content in a enhanced fashion) and broadcast providers (increasingly telcos in „triple or quadruple play" offers). Through industry participation in the consortium, we have first adopters of the technology, a clear target for exploitation and a means for user trials and business model proofing.

In LinkedTV, the public broadcaster **RBB (DE)** will provide broadcast and additional content, tailored to implementing one or more usage scenarios. It will plan and perform user trials and explore the potential for technology- and service-related exploitation for the project. The research experience of RBB includes developing future usage patterns and scenarios, content metadata extraction and personalization, content management and publication, dynamic user services creation and user testing and evaluation

Complementing this with their extensive media archive, **Sound and Vision (NL)** will enrich the LinkedTV scenarios with a rich source of cultural heritage and news media content.

Finally, the POLETI department of the University of Mons (BE) has acquired an expertise in natural interaction approaches, which provides non-intrusive ways of analyzing behavior. This expertise is needed for workpackage 4 and more precisely for task T4.2. Computational models of human attention are also part of the equation, with models focusing on efficiency and real-time processing in ecological situations, with both bottom-up and top-down models implemented on still images, audio and video signals. The variability of the signals which can be processed and the efficiency and robustness of the models on real-life data are two key points needed by the ConnectME project. Moreover, machine learning algorithms and interactive visualization/search have been applied in the framework of multimedia data including music, images, videos or laughter signal. Finally, POLETI/UMONS recently founded, and hosts, the NUMEDIART Institute for New Media Art Technology, providing a connection with the creative arts, and means for creating appealing showcases*.

## *2.4. Resources to be committed*

LinkedTV work will be done with a total of 896 person months over 42 months, representing 21.33  persons' effort per month with 12 partners asking for funding. The total funding requested is ca. 6.76 Mill. €

LinkedTV is an Integrated Project. It is directed towards innovations in one of the central fields of Networked Media where expertise from different domains has to be integrated. We have partners from industry including SME, from public broadcasting organizations, and from research institutions and universities. They all have to provide their own dedicated contributions, and they have tightly to cooperate in order to achieve our goals.

The four WPs 1-4 are directed towards methodological research. They integrate efforts from different domains related to our methodological research. Mainly Fraunhofer, CERTH, UEP, CWI, and EURO contribute their methodological research and experiences here in joint efforts. All four WPs have a comparable workload (with some more emphasis on the annotation and production support side). Especially in WP3 (LinkedTV interface  and presentation engine) the three use case providers and public broadcasters will take an active part in order to make sure that their requirements to these issues are fulfilled.

WP5 is dedicated to the creation of the LinkedTV platform. This is typically a bigger effort including large amounts of software engineering and integration. Our partners Condat and Noterik will take most of the efforts here. These resources will be needed in order to build the integration platform including those efforts to integrate the software modules from WPs 1-4.

WP6 is dedicated to the creation and evaluation of our three use cases. Each of our use case partners RBB, UMONS, and S+V will use their 21-22 PM to prepare the use cases raw material, to define the requirements, specify the scenarios, contribute to their implementation on the LinkedTV platform, and evaluate the fulfilment of their initial requirements. They will be supported by STI which will take the lead for this WP6. This guarantees that all three broadcasters can contribute to the overall approach of LinkedTV in a coordinated way, and that the bridge between use cases, methodological research, and platform creation and integration is strong and active.

WP7 collects our dissemination activities. The amount of dissemination activities corresponds with 60 PM to the need of this Integrated Project. Especially the standardization activities which form a central part of the whole project are contained in this WP. Our partner STI is leader of this WP, and together with EURO which will lead the standardization efforts they will get the largest portion of the work load in the WP. The demonstration activities are

integrated in this WP, too. Because they are tightly related to all our RTD we summarize them under the RTD label.

WP8 contains the exploitation efforts of all in all 77 PM. USG leads this WP with the greatest amount of work load. This is especially dedicated to market analyses and the development of business models and exploitation strategies. Our industrial partners and the broadcasters support these activities with sufficient allocated effort for a meaningful contribution. Additionally, the Fraunhofer IAIS and STI Internationalboth contribute to exploitation as organisations specialised in transferring research into commercialization opportunities and industry uptake.

The management WP9 is deliberately lean. The main effort in this area will be provided by Fraunhofer as IP coordinator. All other partners will get a small amount of work load for their management activities. They will be tightly integrated with all other activities in the IP.

# Section 3. Impact

One central result of LinkedTV will be a framework that will enable rich annotation of audiovisual material with associations to concepts, and on the basis of these annotations related Web content can be associated dynamically with objects in the video stream. An end user receives a video in which objects can be accessed and multimedia presentations received to communicate information about those objects. This framework can form the basis for added value services to users, which generate new value in video content by making it interactive and informative. Here we focus on two aspects of the expected impact of LinkedTV: in society, and in business.

## 3.1. Social and business impact

The Lisbon declaration stated that the EU should become "the most competitive and dynamic knowledge based economy in the world" (2000). Competitiveness and dynamism is based on access to information which is available at our fingertips, and which is intuitive and relevant. This is also the central research goal of LinkedTV. Our focus on industrial and social uptake aims to ensure that we not only achieve our research aims but that this result has a real and viable impact in both industry (strengthening European competitiveness) and society (advancing the information society).

The impact of this is to change our understanding of IP-based audiovisual content and particularly IP-based Television. Television has already established itself as a passive medium, and efforts to promote Interactive Television must fight against this ingrained expectation. The Web, on the other hand, has always been interactive and the convergence of the Web with TV must not turn the Web experience into a passive experience like television viewing.

A new paradigm of IPTV merges the popularity of television as a means of communication with the benefits of the Web as an immense content source. We expect this can impact in different ways: (1) in society, European citizens as part of the information society have eased access to relevant information about objects addressed in video, independent of language or national barriers; (2) in business, professionals can use video more effectively as an information source (the various business and financial TV channels and programmes show already how video is used) because the access to related information about a topic in the program can be more readily available for perusal, being presented in an intuitive and meaningful way; (3) in industry, the growing sectors of TV & video production, TV broadcasting and software firms supporting the development and consumption of broadcast media, challenged by the growth in Web video which is already clearly seen as a competitor of traditional broadcasters, need new services to gain a competitive edge and increase market share. The LinkedTV framework can offer this new added value to bring European industry ahead competitively in the increasingly global market.

To achieve this impact, the relevant communities must be prepared, standards sought to ensure industry uptake and early adopters found to demonstrate the value of LinkedTV services. The consortium has prepared itself well for all these challenges, as outlined beforehand and described in detail in this document.

## 3.1.1. Social: strengthening digital society

Education and access to information are two important pillars of the European digital society of the future. Here, Web based audiovisual material and TV broadcaster archives represent a definitive and significant source for both education and information.

Digitization of analogue audiovisual material will ensure future access. Over the past years, technologies for large-scale migration have matured. The same is also true for thinking about migration projects in terms of their efficiency and the workflow models they could follow. Although the process is far from complete, approximately ten million hours of European audiovisual material has already been digitized [7]. Recently, the audiovisual production process shifted from analogue to digital. This so-called 'born digital' content is directly ingested in asset management systems and will also be kept for posterity as electronic files. Due to digitization and digital production, audiovisual content collections are transforming from archives of analogue material into very large stores of digital data.

For the humanities, digital cultural heritage sources are a fundamental dataset. In the Netherlands, Academia (www.academia.nl), developed by Sound and Vision, offers online access to thousands of hours of streaming video for higher education purposes. Similar services are being created across Europe. As television studies have grown and the study of television history in particular, so has the need for original sources. While historians have long been reluctant to use media for research of the 20th century past, they increasingly use audiovisual sources to fill in gaps, to shed new light on where traditional sources only reveal part of the story [6]. The need for digitized sources along with expertise, standards, tools and services to support use and reuse of content has been acknowledged, but this need is still far from being fully met.

Digitization is also a driver to establish new services. Distribution over networks, interoperability with other collections and flexible integration in other environments are just a few of many properties in this new era of enormous potential for audiovisual archives. Therefore, large-scale digitization efforts do not only ensure long-term access, but also have the potential to reveal the social and economic value of the collections. LinkedTV will allow this material to be exploited in a new way, namely by

1. adding a more fine-grained, concept specific, level to access the collections and
2. combining materials of different types, sources and origins to discover and demonstrate new knowledge about the artifacts.

### 3.1.2. Business: strengthening IPTV in Europe

IP-based video (including IPTV) is a growth market globally. The global IPTV subscriber base is expected to grow from 20 million in 2008 to almost 90 million in 2012. LinkedTV can provide the technologies to place Europe at the innovative forefront of IPTV service creation and provision. Already, Europe is being heralded as the potential market leader in IPTV[29] with 42% of the world's IPTV subscribers and nearly half of the sector's revenues by 2010. In actual numbers, this translates to 21.3 million IPTV users generating $5.1 billion of IPTV service revenues in Western Europe by 2010[30].

Yet television, while playing a significant role in the everyday life of citizens globally, remains a largely passive experience with limited interaction possibilities. The much heralded convergence of television and the Web still tends to mean only that content from both sources share the same cable, and that the content may be switched between or viewed in parallel on the same device. For true interaction with digital content, citizens still need to switch to their computer or other Web-enabled device, relying on a browser or associated plug-ins.

Faced with the explosion of digital audio-visual content on the Web, much of it user-generated or independently produced, the traditional media creators and deliverers – including television

---

[29] Europe to dominate IPTV growth, Light Reading, 5.3.2007
http://www.lightreading.com/document.asp?doc_id=118668
[30] Europe to dominate IPTV growth, Light Reading, 5.3.2007
http://www.lightreading.com/document.asp?doc_id=118668

companies and telecommunication operators - face a challenge to their business models as users face broader choice and greater flexibility in their viewing activity. Despite traditional media players tending to have higher quality, more professional, and branded material , the digital citizen is drawn to innovative new ways to access and interact with digital content, where the traditional marketplace is in danger of finding itself playing catch-up. Established "old" players risk losing touch with the younger generation (especially the under 20s) who have grown up with the Internet and for whom the idea of a shared television viewing experience or the need to watch a particular program at a particular time is something hard to imagine[31].

The investment in IPTV (estimated at a total $21 billion[32]) needs to be complemented with new innovation which can use the convergence of TV and Web content to provide new types of interactive service to the digital citizen. This can help produce new market opportunities, particularly for SMEs, protect the business models of the traditional media companies, give European IPTV technology a competitive advantage in a growth market and support the digital society by enabling new forms of information access to citizens.

If this succeeds, IPTV can become a "potential goldmine" for the key business players in the IPTV market[33]. Service providers, it is found, "to make any headway at all, ...they will need to develop unique and compelling applications". We believe that the converged services that can be made possible over the LinkedTV framework can hold the key to establishing European leadership in IPTV technology and to unlocking the "goldmine" of new compelling IPTV applications.

### 3.1.3. Why Europe?

The work of LinkedTV is necessarily European because (a) it is within Europe where much of the leading research in the relevant fields (video analysis, Web data mining, multimedia annotation, Linked Data, user interfaces, multimedia presentation) is taking place (b) it is within Europe where emerging players in sectors of TV & video production, TV broadcasting and software for the development and consumption of online media are encountering huge opportunities for growth and yet strong global competition, and it is the unique innovation and research skills of European multimedia researchers who can provide new solutions to give them the competitive edge.

### 3.1.4. Why will we succeed?

With regard to other research activities, the research community relevant to LinkedTV is sufficiently small to be aware of other work, and big enough to carry out the proposed work. The consortium is European in nature, with many of the consortium members being key players in their respective research area, equipped with excellent contacts to other actors in the community. Hence, through conferences and other research events, the consortium is able to keep abreast of developments outside of LinkedTV. To formalize this, a specific task has been created to monitor ongoing activities to identify potential developments which may impact the workplan. Through early recognition of such developments, the LinkedTV management structure is in a position and has been assigned the powers to consider and execute workplan changes, e.g. take advantage of external results or refocus areas in the workplan. As part of the dissemination activities, partners have the opportunity to identify potentially related activities

---

[31] Klingler Walter: "Jugendliche und ihre Mediennutzung 1998 bis 2008" (Media usage of young people 1998 to 2008). In: Media Perspektiven 12/2008, pp. 625-634
[32]  IPTV'S Economic Realities, Light Reading, 21.11.2005
http://www.lightreading.com/document.asp?doc_id=84604&WT.svl=deptop_1
[33]  IPTV a potential goldmine, Light Reading, 28.11.2005
http://www.lightreading.com/document.asp?doc_id=84783

at national and international events and reach out to external partners to determine potential complementary activities, or discuss overlaps. However, it can also be clearly stated that, to the best of the knowledge of the consortium, no existing project or other research activity focuses on the central goal of LinkedTV: intuitive access to objects in video to provide dynamically generated multimedia presentations about concepts using media extracted semantically from the Web. Many other activities in the multimedia research community are complementary to tasks or sub-tasks of LinkedTV, and the partners of the consortium are already involved in or aware of such activities (and will continue to do so in the duration of the project) and hence are well suited to guide their workpackages to successful, innovative and new research results. In fact, partners are often themselves involved in other activities, giving LinkedTV strong co-operation opportunities. In particular, standardisation activities will be carefully monitored and to the best of the partner's abilities influenced by our work, so that standards in this area are complementary to our results.

The achievement of this impact is most heavily predicated on the non-emergence of a similar solution in the project duration. This may occur if there is work taking place in the field which is not public, e.g. within a commercial organization. Market and product reports will attempt to anticipate new developments in the area of IP-based video provsion and related sectors of relevance to the LinkedTV work, and ensure LinkedTV's uniqueness in what will be a rapidly growing market. As partners come from the broadcasting, content provision, network operation, server side software and client side software sectors, we are in a good position to keep in touch with developments in each of these areas. Also, each of the partners has a strategic interest in knowing the state_of-the-art in their respective business area. Given the preconditions for the LinkedTV framework, including the multimedia semantic modeling and meeting research challenges such as concept extraction at instance level or generating dynamically multimedia presentations, it is unlikely that a commercial solution will emerge in the short term sufficiently close to the LinkedTV objective (see also the state-of-the-art section for an indication of what can be expected in the next 1-3 years)

## 3.2. Positioning with respect to the realisation of a long term vision in the ICT domain

In the next phase of the Web's development we will see ubiquitous access to audio-visual content over IP (on TV, mobile, other devices). As a result of LinkedTV, this trend will be enriched by the possibility of appealing services in which media is conceptually annotated and hence dynamically aggregated with relevant content from other sources on the Web. This can only be made possible in terms of automation and disambiguity through the link to the proposed Connected Media Layer. This can be extended by user and context-personalised aggregation of relevant associated data, e.g. through location awareness, the concept-related content is further refined to also be location-relevant. We see this as a fundamental part of the emerging information society being promoted within Europe, where every European citizen is part of a digital community in which there is free access to information, independent of language or national barriers.

Our focus in the project will be delivery of Web video-based services to PCs and IPTV programming-centred services to Set Top Boxes, while there is potential here to apply the framework to the ubiquitous Web, which may include mobile devices, home devices and even virtual reality-like environments. LinkedTV services should be accessible in daily private life, and in daily business activities, enriching activities by combining the ease of communication afforded by video with the added interactivity to ensure intuitive information access.

The result of LinkedTV is expected to be a society in which television and Web converge in any device, and where video – already proving online to be extremely popular – is

ubiquitously available in an active, and not just passive manner, so that video consumption in the context of ubiquitous Internet access – in contrast to traditional television – can be an educational and stimulating activity rather than breeding a new form of "online couch potato".

The establishment of the LinkedTV framework as a tool for generating added value services, in which video is annotated and delivered to users in an interactive manner, where objects can be selected and information about them intuitively shown to the user, is the necessary precondition for the growth of such services everywhere where there is Internet access, including IPTV, leading to a paradigm shift in television which brings in the interactivity that made the Web so popular, and avoiding the plague of passivity which has inflicted many television viewers.

For the citizen, LinkedTV will mean getting more out of video, and this will be of particular value in social contexts (health, education) as well as business. For industry, LinkedTV can mean market leadership in a promising growth sector (IP-based video distribution and consumption). For European society, LinkedTV can be the basis for ubiquitous and intuitive access to relevant information through Web-converged TV broadcasting.

### 3.2.1. The Networked Electronic Media vision of 2020

**From today's vision of Networked Media, we will create the basis for the next step - interlinking that media to form meaningful concept-based associations for better information retrieval and presentation - which we call Connected Media**

The strategic research agenda of the NEM initiative foresees a future Media Internet which not only offers the audiovisual services we know today such as television, « *but also a wide range of interactive services across all realms of information, education, and entertainment, offering a wide range of new business opportunities* ». LinkedTV contributes directly to this vision, as LinkedTV services will represent an innovative new way to interact with audiovisual material, have instant access to relevant information and associated media, and build new business models around these interaction and access possibilities.

### 3.2.2. The long term outlook for Future Media Networks

**LinkedTV will create a platform for hypervideo, which can be to the Future Internet of Networked Media as HTML was to the success of hypertext on the World Wide Web**

LinkedTV services represent a new interaction paradigm with audiovisual content which can be seen as the fulfilment of hypervideo[34]. It will be able to provide the platform for interactive services on future content networks, beyond our initial application to Web-based and broadcaster-sourced content over IPTV and PC devices. Combined with future technologies like 3D internet and virtual realities, *the LinkedTV platform will deliver hypervideo content and applications to users of Future Media Networks much as HTML and HTTP form the core technology of today's Web experience.*

---

[34] From the Wikipedia article (checked 18.3.09): "hypervideo combines video with a non linear information structure, allowing a user to make choices based on the content of the video and the user's interests". Later in the article, it is stated: "one can imagine an interlinked web of hypervideo forming in much the same way as the hypertext based World Wide Web has formed. This hypervideo based 'Web of Televisions' or 'TeleWeb' would offer the same browsing and information mining power of the Web, but be more suited to the viewing experience" http://en.wikipedia.org/wiki/Hypervideo

### 3.3. Dissemination and exploitation of project results, and management of intellectual property

### 3.3.1. Dissemination

Dissemination is regarded as a two-way dynamic and interactive process, which will be continuous and progressive. Dissemination will be effected at both the consortium level and at partners' level and will be effected through the following channels:

- Setting up of the project web site with clearly visible and regularly updated project information. The site will be set-up by STI International and present the LinkedTV vision as a 'story' grounded in non-expert, real world, meaningful language. It will keep interested parties informed about project developments and project news. Furthermore, the project website will be the main point of call for external inquiries and serve as the project's "business card".

- Presentation of the project and its results as they become available at related thematic national and international events, workshops and conferences. This activity will be led by STI International, and all partners will actively participate in it.

- Participation in related events organised by the EC, including the IST programme's annual events and relevant concertation meetings. All partners will play an active role here. It will be co-ordinated by STI International.

- Participation in pre-standardization and standardization activities. STI International provides a pre-standardization service.

- Establishment of close links with other FP7 projects active in the same areas as LinkedTV as well as related areas (all partners involved)

- Publications in scientific and industrial journals.

- Production of promotional material (posters, brochures, videos, etc.). The purpose of this activity is to "spread the word" about our aims and activities, and do so in an understandable, easy-to-understand way. It will be led by STI International and supported by all project partners.

- Establishing a page in Wikipedia about LinkedTV and its related technologies

- Establishing a presence in the Social Web to raise public awareness (Facebook, blogs, bookmarking sites etc.)

The target audiences for the foreseen dissemination activities are the following:

- The EC – ICT community: to raise awareness for our objectives and trigger collaborations with projects, initiatives and undertakings sharing similar or complementary goals.

- The scientific community: to spread the scientific results and help trigger their use in other areas as well.

- The broadcasting and telecommunications industry: to promote LinkedTV results in relevant industrial sectors through information days, and promotion at industry events. Also, wider uptake of the technologies by tutorials and workshops for professionals will be supported.

- The general public: Targeted, initially at project launch, to inform peers and actors close to the individual partners from all sort of areas in order to interest them in the project as such and raise awareness in order to gain support. As the project matures, the prime dissemination target will move more to parties active in related sectors

(details above). However, it is a goal of the consortium to continue involving the so-called "general public" as much as possible and deemed suitable. After all, it is taxpayers' money that funds activities such as LinkedTV. Consequently, the consortium sees it as its responsibility to also inform the public who makes such undertakings and activities possible.

A number of deliverables, milestones and events have been planned to assure the effectiveness of the general dissemination activities. Details can be found in the project workplan. Deliverables will be, where appropriate and possible (such as for the research tasks),  research publications whose value will be demonstrated by successful peer review (e.g. submissions to conferences or journals).

The main tool for organising and steering dissemination activities will be laid out in the "Plan for the dissemination and use of knowledge".) The LinkedTV partners are well positioned in the academic community to undertake effective scientific dissemination of results, given both their leading roles in the publication of work and the organisation of events in the relevant fields. Some of the top conferences in which partners publish or (co-)organize, and where LinkedTV work will be disseminated, include:

The project partners will make use of their well-established networks and present the LinkedTV results in the following conferences and workshops:

- GMF - The Global Media Forum primarily highlights the rapid technological development in the world of media and debates the ensuing questions related to modified usage.
- ESWC - The Annual European Semantic Web Conference, sponsored by STI International, presents the latest results in research and applications of Semantic Web technologies
- ESTC - European Semantic Technology Conference organized by STI International
- ISWC - The International Semantic Web Conference is a major international forum where visionary and state-of-the-art research of all aspects of the Semantic Web are presented.
- ASWC - Asian Semantic Web Conference co-chaired by STI International
- FIS - Future Internet Symposium
- BIS – Business Information Systems
  Regularly publish
- WIAMIS - International Workshop on Image Analysis for Multimedia Interactive Services
- SAMT - International Conference on Semantics and Digital Media Technologies
- CBMI - International Workshop on Content-Based Multimedia Indexing
- ICMR - ACM International Conference on Multimedia Retrieval
- ICIP - IEEE International Conference on Image Processing
- ICME - IEEE International Conference on Multimedia & Expo
- EUSIPCO - European Signal Processing Conference
- IEEE ICSC - IEEE International Conference on Semantic Computing
- Museums and the Web
- International Federation of Television Archives (FIAT-IFTA) annual conference
- International Broadcasting Convention (IBC)
- Internationale Funkausstellung (IFA)
- European Conference on Digital Libraries
- Relevant events organised by EBU, Europeana, Open Video Alliance
- ACM Multimedia

It is planned that LinkedTV results will be promoted at these leading conferences and play a central role in the co-organized events, which already represent key activities in the field of semantic multimedia systems. It is furthermore foreseen that invited speakers will address

relevant topics of the call, conduct workshops focused on LinkedTV themes, and carry out system demonstrations and/or promotional events as part of the conference activities. Furthermore, and as indicated above, partners already publish in leading journals in the field and are expected to publish project results there, too:

Some journals / publishers that will be targeted include:

- IEEE Transactions on Circuits and Systems for Video Technology,
- IEEE Transactions on Multimedia
- IEEE Transactions on Image Processing
- IEEE Signal Processing Letters
- Computer Vision and Image Understanding, Elsevier
- Signal Processing: Image Communication, Elsevier,
- Journal on Advances in Signal Processing, Hindawi,
- Multimedia Tools and Applications Journal, Springer
- Information Processing & Management (Elsevier)
- Journal of Digital Information (University of Texas)
- Ariadne (UKOLN)
- D-LIB (Corporation for National Research Initiatives)

Other FP7 projects are also a potential target of dissemination activities, as well as further co-operations and complementary activities (e.g. joint events). Many partners in LinkedTV are involved in relevant FP7 projects indicating co-operation and cross-project dissemination opportunities. Besides their partnership in LinkedTV, consortium members are also involved in following related R&D projects, among others:

- CASAM: Computer-Aided Semantic Annotation of Multimedia (2011)
- NoTube: Networks and Ontologies for the Transformation and Unification of Broadcasting and the Internet (2012)
- GLOCAL: Event-based Retrieval of Networked Media (2009-2012)
- CHORUS+: Coordinated approacH to the EurOpean effoRt on aUdio-visual Search engines + (2010-2012)

Standardization efforts will be pushed through active involvement in the W3C standardisation activities, such as the W3C Multimedia Semantics Incubator Group and the W3C Multimedia Annotation on the Semantic Web Task Force. LinkedTV's advancement is expected to foster ongoing efforts by consortium members of applying Semantic Web technologies in Multimedia Semantics within the above standardisation activities. Hence we expect LinkedTV will have substantial impact on further standardization efforts of:

LinkedTV partners are uniquely positioned to contribute due to their past and existing activities and contributions in the area:

CERTH-ITI actively participates in standardisation activities. In the near past it participated in activities such as the W3C Multimedia Semantics Incubator Group, http://www.w3.org/2005/Incubator/mmsem/, which among others has published the "Image annotation on the Semantic Web Incubator Group Report" and the "Multimedia Vocabularies on the Semantic Web Incubator Group Report". CERTH-ITI also participated and continues to partcipate in ISO/IEC JTC1/SC29/WG1 (JPEG) and ISO/IEC JTC1/SC29/WG11 (MPEG) standards working groups, and as part of its work it has authored and submitted several input documents to these working groups' meetings.

The W3C has recently stated its strategy to make video a first class citizen on the web. A new Video Activity, following the Multimedia Semantics Incubator Group co-chaired by Raphaël Troncy, has been launched encompassing several working groups. Among others, the W3C Media Fragments Working Group  will address temporal and spatial media fragments on the web using Uniform Resource Identifiers (URI). URI-based addressing of spatio-temporal fragments of audiovisual news content on the web will be a key issue for LinkedTV. Having

global identifiers for clips and media objects would allow substantial benefits, including linking, bookmarking, caching and indexing the content. Furthermore, the W3C Media Annotations Working Group will develop an ontology designed to facilitate cross-community data integration of information related to media objects on the Web. This ontology will be compatible with the knowledge infrastructure put in place in LinkedTV. Consortium member Eurecom will actively participate and co-chair these working groups.

Also, project results will be fed into the activities of the European Broadcasting Union (EBU), especially with regards to interactive services and metadata provision. We will also outreach to The Open Video Alliance which is an ad-hoc group of organizations dedicated to fostering the growth of open video.

Naturally, and in addition to what has already been outlined above, dissemination activities will also be undertaken with the clear goal to address the markets targeted for potential exploitation: broadcasters, content owners, telcos, IPTV network operators and platform providers, as well as equipment manufacturers (STBs).

Consortium marketing departments will assist in the design of business oriented versions of promotional material, and specific media opportunities will be identified and undertaken by the media partners in the consortium.

### 3.3.2. Exploitation by partners

It is planned to pursue a two-way exploitation strategy in LinkedTV:
- On a joint level (the consortium as a whole or a combination of consortium partners)
- On an individual partner level.

Joined exploitation strategies will be developed for LinkedTV results that are being developed collaboratively. Details of joined exploitation strategies will be agreed in the Consortium Agreement and in the Exploitation Plans and Agreements to be developed. This will be done as project results are becoming available and will be updated regularly. In addition to exploitation being carried out jointly on a collaboratively level (e.g. of), individual partners will also pursue exploitation activities individually. Besides possible new products and services, this includes, for example, things such as changed business practices, improved workflows, partnerships established with the help of LinkedTV work and activities, and the like.

The following sections will first provide an overview of major market trends and then more information on the way the consortium plans to address exploitation issues both on an overall project level, as well as on an individual partner level.

Regarding commercialisation of the LinkedTV framework, it is foreseen that users subscribe to particular services which – potentially independent of the AV content or its annotation – provide particular types of data aggregated to the AV content. Besides entertainment and monetisation opportunities (e.g. companies provide sponsored services which promote their products), other use cases are
- education
- academic or commercial research
- public information dissemination

Exploitation will be led by the industry members of the LinkedTV consortium. Following this, aided by the early adopter effect, we expect exploitation of results to be able to be spread to external companies.

Firstly, we look at market trends and exploitation potential in general:

### 3.3.2.1. Market trends

IP-based video (including IPTV) is a growth market globally. The global IPTV subscriber base is expected to grow from 32 million in 2010 to almost 120 million in 2014. Global IPTV service revenues are forecasted to grow from $11 billion in 2010 to nearly $50 billion in 2014. On average, global IPTV subscriber base and global IPTV service revenues will grow between 35% and 50% per year between 2010 and 2014[35]. Other research is more conservative predicting a weak economic atmosphere through 2011 and therefore revising the 3-year average annual growth rate of IPTV subscriber base from 29% to less than 15%[36].

Europe is being heralded as the potential market leader in IPTV[37] with 40% of the world's IPTV subscribers and nearly half of the sector's revenues by 2010. In actual numbers, this translates to 34 million IPTV users generating $9.1 billion of IPTV service revenues in Western Europe by 2012[38].

According to most recent data, Western Europe currently has 14 million IPTV subscribers. This is more than two times of the respective subscriber base in North America and still ahead of Asia. However, while Western European subscriber base has increased by 50% from 2008 to 2010, it has doubled in North America and South/East Asia[39]. The faster growth of the subscriber base in Europe provides an advantage for European providers, but at the same time creates higher pressure for additional services that go beyond current offerings including linear TV, video on demand and exclusive content. European providers have the possibility for a first move advantage in sophisticated IPTV services. To keep the leading role in the IPTV market, European IPTV operators must now leverage the characteristics of the new platform to produce a differentiated offering, redefining the experience of television." Television remains a highly popular medium for education and entertainment. Despite the growth in the consumer electronics market, the television set is likely to retain its central place in our living rooms (and other rooms..) in the future. Hence, many consumer electronics companies choose to focus more on shifting functions of other devices into our television sets. According to major market research companies, the main options how to improve the IPTV offerings are[40]:

- Integration of Web content (text or video) with traditional broadcast content.

- Extending the offerings to any Internet-connected device, e.g. a PC, mobile handset, or portable device.

- Targeted advertising as IPTV's interactivity and personalization allow much more effective ad strategies than conventional broadcast.

Interactive, sophisticated and personalized services, could add between 15-22% of telecom operators' retail revenue in the top 5 Western European markets. In the UK alone, it is expected that TV revenues could grow at an 8% 5-year CAGR[41].

---

[35] IPTV Global Forecast –2010 to 2014, Multimedia Research Group, October 2010, http://www.mrgco.com/iptv/gf1210.html

[36] Research in Focus - IPTV Offers Analysis of IPTV Subscription Trends, As Well As ARPS and Revenue Trends for the 2006-2015 Period www.tvover.net/2010/12/14/Report+Research+In+Focus+IPTV+Offers+Analysis+Of+IPTV+Subscription+Trends+As+Well+As+ARPS+And+Revenue+Trends+For+The+20062015+Period.aspx

[37] IPTV Leads the Telco Transformation in 2009, TVover.net, 26.2.2009 http://www.tvover.net/2009/02/26/IPTV+Leads+The+Telco+Transformation+In+2009.aspx

[38] IPTV Potential for Europe, www.lightreading.com/document.asp?doc_id=199901&site=lr_cable

[39] GOLDMEDIA TRENDMONITOR 2011, www.goldmedia.com/aktuelles/trendmonitor-2011.html

[40] www.abiresearch.com/press/1127-IPTV+Operators+Searching+for+Competitive+Differentiation)

[41] (www.tvover.net/2009/02/26/IPTV+Leads+The+Telco+Transformation+In+2009.aspx

### 3.3.2.2. Exploitation Strategies of LinkedTV Partners

LinkedTV will develop cutting-edge technology that enables interactive connection of AV and Web content. The LinkedTV technology will provide the potential of first mover advantage of the companies involved in the project and other European companies in the industry.

At present, telecommunications companies base new business models around delivery of TV over IP (IPTV) bundled with other services made possible by the integration of the Internet/Web in the broadcast network. IPTV broadcasts are generally no different from TV broadcasts over other channels, the „IP" only introduces a means to offer television, telephony and Internet access in one line („Triple Play") and deliver other applications to the television (via the Set Top Box) which exist alongside the normal TV broadcast option.

LinkedTV, on the other hand, will deliver new technology which will allow IPTV broadcasts to be radically different from traditional TV broadcast through the integration of Web technology. Telcos offering services based around LinkedTV can expect to profit from the added value of such services over their competitors. For such services to exist, the media providers will be enabled by LinkedTV to enhance their broadcasts in dynamic, intelligent and customizable fashion. Through the LinkedTV-enabled analysis and annotation of broadcast material, the set of LinkedTV technologies can be used to create added value services without prohibitive manual preparation. Media providers can increase market share by offering their content over LinkedTV services, winning viewers through the new possibilities enabled to interact with the content. The different players in the market – media owner, producer, broadcaster, distributer – may each have specific types of service developed which reflect their intended aims with the content, whether it be education, gaining feedback, building communities, advertising or winning market share.

Through LinkedTV, interaction with television will reach maturity as usage will be based on the content of broadcasts – hence tightly integrated with the viewing context. New and innovative opportunities will arise for stakeholders of different types to reach viewers – and achieve different aims, from education and social cohesiveness to marketing and higher profits.

Against the background of the overall market trends and the specific opportunities created by LinkedTV technology, the following section highlights exploitation plans on a project partner level.

**CERTH-ITI** is already participating in the VR-Sense spin-off company with the aim to produce and distribute innovative high technology products based on research results and results from R&D projects. Furthermore, part of the CERTH-ITI business plan is to participate in a number of new spin-off commercial companies capable of exploiting its research when new market needs and solutions are identified. The Centre for Research and Technology Hellas (CERTH), where ITI belongs, was in fact founded so as to accomplish Technology Transfer, encouragement of entrepreneurship and innovation. The Informatics and Telematics Institute as a member of CERTH has all the necessary support including legal support, business management, marketing, distribution sales channels and accounting in order to create innovative enterprises. CERTH-ITI is active in providing research services and results to the local and European industry through direct research contracts and licensing agreements.

Furthermore, depending on the progress of the industrial exploitation efforts, a suitable subset of the techniques implemented by CERTH will be made available in the form of open source software, under an appropriate open-source license and according to the exploitation plan decided during the project. This will enable its use outside the consortium. As part of its commitment to open source software, CERTH has made available in the near past the source code of tools developed by it, such as the VIA: Video and Image Annotation Authoring tool

(*http://mklab.iti.gr/via/*) and the GPU-accelerated implementation of LIBSVM (*http://mklab.iti.gr/content/gpu-accelerated-libsvm*). The foreseen open-source distribution will facilitate the further exploitation of the project's results and allow CERTH's technology to be integrated with other third party tools.

**CONDAT** expects to exploit the project results by reusing methodology and software components developed within the project to enhance their existing TV- and media solutions by the advanced annotations to browse and retrieve videos. Condat´s customers in the broadcasting and Web-TV sector have expressed growing interest for such features. The main strategic impact on Condat's business is to lay foundations for a competitive advantage for their Condat® Media Suite. This solution comprises modules for planning, media asset management and editing for different distribution channels such as DVB-T, DVB-H, Web-TV or videotext. The functional modules of the Media Suite range from 30.000 up to 250.000 €. Condat's market focus is medium to large enterprises, especially in the TV, media and telecommunications sector.

The new features of LinkedTV will allow the customer to offer interweaved content, better search possibilities and browsing which makes the whole presentation more attractive and optimises working time for the end users. This allows Condat to raise the prices for product and support, especially because of the advantage related to other products in this market segment. The additional revenues planned will reach the ROI in a few years related to the investment in the project. The consortium agreements concerning IPR and exploitation will allow Condat to enhance their existing solutions with the main features developed in the project.

**Sound and Vision** intents to implement the technology provided by LinkedTV to its current online portal. Sound and Vision aim to introduce hyperlinked video as an integral part of its online services Additional resources will be invested to implement this knowledge in the current catalogues as ORPHEUS releases the technology.

Furthermore, Sound and Vision positions itself as one of the key players in the development of Europeana (part of EDLnet), through its involvement in the Video Active project, and as a research organization highly active in the areas preservation and digital durability (undertaken for example in the PrestoPRIME project). It is also expected that LinkedTV technology will be used to enhance access to material on Europeana.

Finally, Sound and Vision is also an active participant in the evolving landscape of European media production, working closely with the EBU and the CHORUS Coordinated Action on standardization issues. Fruitful co-operations and exchanges for mutual benefit are expected on these levels, too.

**Noterik** will implement LinkedTV technology on its video platform. All partners and relations will be notified if new technology comes available to them and so will LinkedTV technology be offerered to a wide range of dutch institutes and companies.

Noterik will also provide LinkedTV technology to possible new clients and will venture into possibilities of exploiting specific LinkedTV technology to a wider range of suitable clients and prospects in the Netherlands. It will also try to exploit this technology within the realms of the EUScreen project.

Primarily, **RBB** will exploit the LinkedTV technology and services internally by integrating the developed environment to enrich its content offerings. Initially, it is foreseen that the online offer will be the prime beneficiary. Later, and once the service operates robustly, it is expected to become more TV-centric. By integrating the LinkedTV technology in its workflow, and by offering it to the content production staff, RBB will be enabled to create enhanced and improved services with no significantly increased production time. This equals

either cost savings or being in a position to produce more / better services with the same resources. If successful, both the system as well as experiences gained can be transferred to other companies / organizations with similar aims and ambitions.

Clearly, on the production side, RBB will benefit from the open architecture allowing easy, seamless integration of content from any source into these new user-oriented meta-services and corresponding new workflows. RBB intends to apply the LinkedTV project results to its production and provisioning of services, programmes and content in order to offer its existing and new audiences the type of services and content they want in a manner that is fair and acceptable to all parties. As a member of the ARD – the Association of Public Service Broadcasting Corporations in Germany, RBB will use its contacts with other broadcasters within the ARD and beyond to promote LinkedTV results and developments using internal events, meetings and workshops. RBB will present the project developments at international exhibitions such as the IFA and IBC, also within the ARD group and the European Broadcasting Union, e.g. the annual Eurovison TV Summit. RBB will also support publications in journals like the EBU Technical "tech-i" and provide input for promotional flyers, brochures etc.

In addition to the exploitation plans outlined above, RBB intends to participate in collaborative exploitation activities with other project partners. Being a co-developer of the technology (especially contributing evaluations, tests and trials) detailed agreements are to be entered into once tangible results materialize. However, being a non-profit public service broadcaster, RBB will not play a major role when it comes to pursuing commercial exploitation activities as, simply put, the required expertise is lacking. However, this said, RBB can well imagine establishing closer partnerships with other consortium members on a separate legal level (i.e. form a new legal entity - something RBB is entitled to by its statutes) in order to deal with the full-scale exploitation of project achievements.

LinkedTV offers new exploitation opportunities for universities and research institutions as well. First of all new, exclusive knowledge related to new technologies, interactive content formats and innovative business models will be developed. Technical universities will explore the opportunities to leverage developed technological components and knowledge through spin-offs or by offering consulting services on how to integrate and develop interactive IPTV and Web content services. The business administration University of St. Gallen will leverage the new knowledge regarding interactive content formats, advertising models and business models to offer dedicated services to the media and telecommunication industry in Europe. The University of St. Gallen will furthermore develop specific executive education courses for the media and telecommunication industry.

Finally, let us consider exploitation potential in a wider industry context:

**Table 9: Exploitation measures**

| Potential Exploitation Partner | LinkedTV Result for Exploitation | Measures to Ensure Exploitation |
|---|---|---|
| TV/Video Production | Tools for annotation | Open source tools, further commercial development, standardisation of annotation schemes, demonstration events |
| TV/Video Broadcasting | Tools for delivery of content-aggregated video streams | Open source tools, further commercial development, standardization of delivery format, demonstration events |
| Software/Services for TV/Video creation & management | LinkedTV platform for annotation & content aggregation | Further development, documentation, licensing, monetarization opportunities |

| | | (server-side advertising) |
|---|---|---|
| Software/Services for TV/Video playback | LinkedTV hypervideo player | Further development, documentation, licensing, monetarization opportunities (client-side advertising) |
| The European citizen | LinkedTV services | Public promotion (e.g. online showcases), user trials, private service offers |

### 3.3.2.3. Portability of LinkedTV technology to other markets

The state-of-the-art technology which will be developed in the course of the project is ambitious and will furthermore generate interactive technology which can be used for innovative new IP-based content services

Several components will be developed based on open standard technology, which will have a high level of interoperability in order to market these specific LinkedTV components to key market players.

Therefore, LinkedTV should not only be exploited as a Web- and STB-based application. In the medium-term, it should find suitable key market players in other network markets who have a need for cutting edge technology from which they can benefit directly. For example, liner TV is at present ported to mobile devices. Most of the European countries have introduced linear mobile TV. In the future, we fully expect mobile video to have a similar development path as IPTV and Web TV and become an early adopter market of these technologies. In particular, as early experiences with mobile TV show that more sophisticated and personalized services that go beyond linear mobile TV are required. No services comparable to LinkedTV are yet available on any of these markets.

The exploitation and market survey deliverables of WP8, will expand on this analysis, identifying and targeting key players in the most promising market sectors.

### 3.3.3. Knowledge Management

Knowledge management is very important in an innovative project such as LinkedTV. Both the *background knowledge* brought to the project by the partners to do the work and the *foreground knowledge* being generated by the work done in the project must and will be clearly identified and appropriately protected. All consortium members agree:

- to bring all necessary knowledge for the performing of their work as foreseen in the workplan to the project without restriction or charge;
- to make knowledge available to other partners where that is necessary for the correct understanding of work done or the successful alignment or integration of separate activities without restriction or charge, other than that the partner may request that said knowledge may not be shared outside of the project;
- where knowledge brought to the project is already (prior to the project) restricted in some way, the partner responsible for providing that knowledge commits to making that knowledge available within the project to the necessary partners in a fair and free manner, possibly associated to restrictions in use;
- commercial partners with IPR may retain those rights, while making available knowledge in the project as is necessary for fulfilling the workplan, or may choose to grant rights to the research partners so that they may carry out the work;
- to make available all knowledge generated as a result of the work done in the project to the project without restriction or charge, other than that the partner may request that said knowledge may not be made public outside of the project;

- where knowledge generated within the project applies to work which already is restricted in some prior and agreed way, that new knowledge will be made available within the project in a manner no more restrictive than the prior work.

In short, while the consortium members agree to a fair and free access to knowledge within the project (this will ensure that no difficulties in carrying out the workplan occur), we wish to ensure that both academic and commercial partners who will have the opportunity in LinkedTV to advance their own research and development, have a basis for protecting the knowledge gained for both future dissemination and exploitation.

While LinkedTV will carry out public dissemination and exploitation activities, some of the knowledge may be protected in a fair manner agreed by the consortium. For example, research advances may be licensed to individual partners, while generally we expect the use of open source and free of charge licenses. This ensures that intellectual property remains the property of the originator but that other research may be able to make use of results to achieve further advances.

Commercial advances should not be restricted by the consortium, especially as this is a major goal of the project. The LinkedTV platform will carry an open source license, so that further development by commercial organisations is supported. We will allow the software partners Condat and Noterik to enhance their existing solutions with the main features developed in the project. The exploitation partners RBB and Sound & Vision will also continue to have access to the platforms for LinkedTV services developed in the project.

The consortium agreement will be the basis in LinkedTV for the clarification of IPR brought to the project or created in the project, and the rights to access for the other project partners and for external organisations, both during and after the project duration.

## Section 4. Ethical Issues

LinkedTV will follow the guidelines (expressed in form of "opinions") of the European group on ethics in science and new technologies to the European Commission. All national legal and ethical requirements of the Member States where the research is performed will be fulfilled.

### *ETHICAL ISSUES TABLE*

|  | YES | NO |
|---|---|---|
| **Informed Consent** | | |
| • Does the proposal involve children? | | X |
| • Does the proposal involve patients or persons not able to give consent? | | X |
| • Does the proposal involve adult healthy volunteers? | | X |
| • Does the proposal involve Human Genetic Material? | | X |
| • Does the proposal involve Human biological samples? | | X |
| • Does the proposal involve Human data collection? | | X |
| **Research on Human embryo/foetus** | | |
| • Does the proposal involve Human Embryos? | | X |
| • Does the proposal involve Human Foetal Tissue / Cells? | | X |
| • Does the proposal involve Human Embryonic Stem Cells? | | X |
| **Privacy** | | |
| • Does the proposal involve processing of genetic information or personal data (eg. health, sexual lifestyle, ethnicity, political opinion, religious or philosophical conviction) | | X |
| • Does the proposal involve tracking the location or observation of people? | | X |
| **Research on Animals** | | |
| • Does the proposal involve research on animals? | | X |
| • Are those animals transgenic small laboratory animals? | | X |
| • Are those animals transgenic farm animals? | | X |
| • Are those animals cloned farm animals? | | X |
| • Are those animals non-human primates? | | X |
| **Research Involving Developing Countries** | | |
| • Use of local resources (genetic, animal, plant etc) | | X |
| • Benefit to local community (capacity building i.e. access to healthcare, education etc) | | X |
| **Dual Use** | | |
| • Research having direct military application | | X |
| • Research having the potential for terrorist abuse | | X |
| **ICT Implants** | | |
| • Does the proposal involve clinical trials of ICT implants? | | X |
| **I CONFIRM THAT NONE OF THE ABOVE ISSUES APPLY TO MY PROPOSAL** | X | |

# References

1. Minelli, S. (2007). Gathering requirements for multilingual search of audiovisual material in cultural heritage. Proc. of Workshop on User Centricity – state of the art (16th IST Mobile and Wireless Communications Summit), Budapest, Hungary.

2. Oomen, J. and V. Tzouvaras (2007), Providing Access to European Television Heritage. In Ariadne. Issue 53 http://www.ariadne.ac.uk/issue53/ooman-tzouvaras/

3. Bates, M. et al. (1993). An Analysis of Search Terminology Used by Humanities Scholars: The Getty Online Searching Project Report Number 1. The Library Quarterly 63(1): 1-39.

4. Dalton, M. S.  and Charnigo, L. (2004). Historians and their information sources. College & research libraries 65:400-425.

5. Wright, Richard (2008) Preservation of Digital Audiovisual Content: briefing paper. Accesed online at:
   www.digitalpreservationeurope.eu/publications/briefs/audiovisual_v3.pdf

6. Nemrava J., Buitelaar P., Simou N., Sadlier D., Svatek V., Declerck T., Cobet A., Sikora T., O'Connor N., Tzouvaras V., Zeiner H, Petrak J.:An Architecture for Mining Resources Complementary to Audio-Visual Streams. In: Workshop on Knowledge Acquisition from Multimedia Content (KAMC) at SAMT 2007, Genova.

7. K-Space NoE, http://www.k-space.eu/

8. Boemie, http://www.boemie.org/

9. N. Sebe, M. S. Lew, X. Zhou, T. S. Huang, E. M. Bakker, "The state of the art in image and video retrieval", in Proceedings of the 2nd International Conference on Image and Video Retrieval, Urbana, July 2003

10. C.G.M. Snoek, M. Worring, "Multimodal Video Indexing, a Review of the State-of-the-art," Multimedia Tools and Applications, Springer Netherlands, pp. 5-35, January 2005.

11. A. Ghoshal, P. Ircing, and S. Khudanpur, "Hidden Markov models for automatic annotation and content-based retrieval of images and video," in Proc. SIGIR 2003, 2003, pp. 544–551.

12. K. Barnard, P. Duygulu, D. Forsyth, N. de Freitas, D. M. Blei, and M. I. Jordan, "Matching words and pictures," J. Mach. Learn. Res., vol. 3, pp. 1107–1135, 2003.

13. C. Dance, J. Willamowski, L. Fan, C. Bray, and G. Csurka, "Visual categorization with bags of keypoints," in Proc. ECCV Int. Workshop on Statistical Learning in Computer Vision, Prague, Czech Republic, 2004.

14. Y.-G. Jiang, W.-L. Zhao, and C.-W. Ngo, "Exploring semantic concept using local invariant features," in Asia-Pacific Workshop on Visual Information Processing, Beijing, China, 2006.

15. N. Voisine, S. Dasiopoulou, V. Mezaris, E. Spyrou, Th. Athanasiadis, I. Kompatsiaris, Y. Avrithis, and M. G. Strintzis, "Knowledge-assisted video analysis using a genetic algorithm," in Proc. 6th Int. Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2005), 2005.

16. Shih-Fu Chang, Manmatha, R., Tat-Seng Chua "Combining text and audio-visual features in video indexing" in Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05), March 2005, Philadelphia, USA

17. J. Fan, Y. Gao, and H. Luo, "Multi-level annotation of natural scenes using dominant image components and semantic concepts," in Proc. ACM Multimedia, 2004, pp. 540–547.

18. Jane Hunter. Adding Multimedia to the Semantic Web - Building an MPEG-7 Ontology. In 1st International Semantic Web Working Symposium (ISWC), pages 261–281, 2001.

19. Chrisa Tsinaraki, Panagiotis Polydoros, and Stavros Christodoulakis. Interoperability support for Ontology-based Video Retrieval Applications. In 3rd International Conference on Image and Video Retrieval (CIVR), pages 582–591, 2004.

20. Roberto Garcia and Oscar Celma. Semantic Integration and Retrieval of Multimedia Metadata. In 5th International Workshop on Knowledge Markup and Semantic Annotation, pages 69–80, 2005.

21. Richard Arndt, Rapha¨el Troncy, Steffen Staab, Lynda Hardman, and Miroslav Vacura. COMM: Designing a Well-Founded Multimedia Ontology for the Web. In 6th International Semantic Web Conference (ISWC), 2007.

22. Raphaël Troncy, Óscar Celma, Suzanne Little, Roberto García and Chrisa Tsinaraki. MPEG-7 based Multimedia Ontologies: Interoperability Support or Interoperability Issue?. In Workshop on Multimedia Annotation and Retrieval enabled by Shared Ontologies (MAReSO'07), Genova, Italy, December 5, 2007.

23. Stephan Bloehdorn, Kosmas Petridis, Carsten Saathoff, Nikos Simou, Vassilis Tzouvaras, Yannis Avrithis, Siegfried Handschuh, Yiannis Kompatsiaris, Steffen Staab, and Michael G. Strintzis. Semantic Annotation of Images and Videos for Multimedia Analysis. In 2nd European Semantic Web Conference (ESWC), pages 592–607, 2005.

24. N Maillot, M. Thonnat and A. Boucher. Towards ontology-based cognitive vision. Machine Vision and Applications MVA, 16(1),pp33-40, Springer-Verlag Heidelberg, December, 2004

25. http://www.metadataworkinggroup.com/

26. http://www.w3.org/2008/WebVideo/Annotations/

27. L. Hardman. Modelling and Authoring Hypermedia Documents. PhD thesis, CWI, Amsterdam, 1998.

28. L. Hardman, D. C. A. Bulterman, and G. van Rossum. Links in hypermedia: the requirement for context. In HYPERTEXT'93: 4th ACM Conference on Hypertext and Hypermedia, pages 183–191, Seattle, Washington, USA, 1993.

29. Z. Li, H. Davis, and W. Hall. Hypermedia Links and Information Retrieval. 14th Information Retrieval Colloqium, Lancaster University, http://eprints.ecs.soton.ac.uk/772/, 1992.

30. L. Hardman, J. van Ossenbruggen, K. S. Mullender, L. Rutledge, and D. C. A. Bulterman. Do you have the time? Composition and linking in time-based hypermedia. In HYPERTEXT'99: 10th ACM Conference on Hypertext and Hypermedia, pages 189–196, New York, NY, USA, 1999.

31. Synchronized Multimedia Integration Language (SMIL 3.0). W3C Recommendation, 01 December 2008, http://www.w3.org/TR/SMIL/

32. S. Pfeiffer and C. Parker. Specifying time intervals in URI queries and fragments of time-based Web resources, IETF Draft, http://www.annodex.net/TR/URI_fragments.html

33. http://youtube.com/watch?v=UxnopxbOdic

34. Dumais, S. T. Cutrell E., Cadiz J. J., Jancke G., Sarin R. and Robbins D. C. (2003). Stuff I've Seen: A system for personal information retrieval and re-use. In Proceedings of SIGIR 2003.

35. Sieg A.,Mobasher B., Burke R. (2007). Ontological User Profiles as the Context Model in Web Search. In Proceedings of the 2007 DePaul CTI Research Symposium, Chicago, IL, May 2007.

36. Sieg A.,Mobasher B., Burke R. (2007) Learning Ontology-Based User Profiles: A Semantic Approach to Personalized Web Search. In IEEE Intelligent Informatics Bulletin, vol. 8, NO. 1, November 2007

37. G. Th. Papadopoulos, A. Briassouli, V. Mezaris, I. Kompatsiaris, M. G. Strintzis, "Statistical Motion Information Extraction and Representation for Semantic Video Analysis", IEEE Transactions on Circuits and Systems for Video Technology, vol. 19, no. 10, pp. 1513-1528, October 2009.

38. V. Mezaris, A. Dimou, I. Kompatsiaris, "On the use of feature tracks for dynamic concept detection in video", Proc. IEEE International Conference on Image Processing (ICIP 2010), Hong Kong, China, September 2010, pp. 4697-4700.

39. G. Th. Papadopoulos, V. Mezaris, I. Kompatsiaris, M. G. Strintzis, "Combining Multimodal and Temporal Contextual Information for Semantic Video Analysis", Proc. IEEE International Conference on Image Processing (ICIP 2009), Cairo, Egypt, November 2009, pp. 4325-4328.

40. V. Mezaris, S. Gidaros, G. Th. Papadopoulos, W. Kasper, J. Steffen, R. Ordelman, M. Huijbregts, F. de Jong, I. Kompatsiaris, M. G. Strintzis, "A system for the semantic multi-modal analysis of news audio-visual content", EURASIP Journal on Advances in Signal Processing, vol. 2010, 2010.

41. N. Gkalelis, V. Mezaris, I. Kompatsiaris, "A joint content-event model for event-centric multimedia indexing", Proc. Fourth IEEE International Conference on Semantic Computing (ICSC 2010), Pittsburgh, PA, USA, September 2010, pp. 79-84.

42. A. Scherp, R. Jain, M. Kankanhalli, V. Mezaris, "Modeling, Detecting, and Processing Events in Multimedia", Proc. ACM Multimedia 2010, Firenze, Italy, October 2010, pp. 1739-1740.

43. D. Tsatsou, F. Menemenis, I. Kompatsiaris and P. C. Davis. "A Semantic Framework for Personalized Ad Recommendation based on Advanced Textual Analysis", 3rd ACM Conference on Recommender Systems (RecSys'09), pp.217-220, NY, USA, October 2009.\]

44. D. Tsatsou, S. Papadopoulos, I. Kompatsiaris, P.C. Davis, "Distributed Technologies for Personalized Advertisement Delivery", in Online Multimedia Advertising: Techniques and Technologies, Xian-Sheng Hua, Tao Mei, Alan Hanjalic (editors), IGI Global, 2010 (to appear)

45. S. Stefanov, Svetlin, V. Huang. "A Semantic Web Based System for Context Metadata Management", in Metadata and Semantic Research, Communications in Computer and Information Science, Volume 46. ISBN 978-3-642-04589-9.
Springer Berlin Heidelberg, pp. 118-129, January 2009.

## Appendix A: Related projects (full list)

| Project description | Similarities and differences with LinkedTV |
|---|---|
| **GLOCAL** (FP7-CP) - Event-based Retrieval of Networked Media. GLOCAL offers a Semantic Web based platform for multi-media retrieval which uses events as the primary means to organize and index media, e.g., photos, videos, journal articles. A local dimension enables the mapping of tags (conceptualizations) to media (personal experiences), while a global dimension enables the sharing of event descriptions thus enabling social sharing and networking of events, across similar events, thus providing a common way to index media. | GLOCAL also applies conceptualization to index and retrieve multi-media materials, but 1) is organized on the basis of events and 2) videos remain again weakly integrated as already explained in the introduction of the LinkedTV proposal. Objects and scenes in videos can not be identified, connected to context and retrieved. LinkedTV will provide this missing gaps to identify objects in videos, link them to different types of context, retrieve them and demonstrate their use in a hypermedia player. As LinkedTV deploys existing standards such as HTML5, MPEG7, MPEG21 and SMIL it would be well suited to enhance platforms such as GLOCAL for video materials. |
| **ALICANTE** (FP7-CP) — Media Ecosystem Deployment Through Ubiquitous Content-Aware Network Environments. Alicante allows users to access multimedia services in various contexts, and to share and deliver his own audiovisual content dynamically, seamlessly, and transparently to other users. This is realised by providing an open, metadata-driven, interoperable middleware for the adaptation of distributed media resources to the user's preferences and heterogeneous contexts enabling an increased quality of experience. | ALICANTE also uses meta data, context and user preferences to retrieve all kinds of multi-media materials, but videos remain again weakly integrated. LinkedTV provides all functions for the end-to-end process to retrieve concepts, to weave materials with other resources from the Web and a hypermedia player to allow new user experiences. The objects are identified not only on a frame level, but also for scenes and video sequences. As LinkedTV deploys existing standards such as HTML5, MPEG7, MPEG21 and SMIL it would allow to enhance platforms such as ALICANTE for video materials. |
| **K-Space** (FP6-NoE) — K-Space focuses on creating tools and methodologies for low-level signal processing, object segmentation, audio processing, text analysis, and audiovisual content structuring and description. It builds a multimedia ontology infrastructure, analyses and enhances knowledge acquisition from multimedia content, knowledge-assisted multimedia analysis, context-based multimedia mining and intelligent exploitation of user relevance feedback. In this process it creates a knowledge representation for multimedia, distributed semantic management of multimedia data, semantics-based interaction with multimedia and multimodal media analysis. http://www.k-space.eu/ | LinkedTV will build on the results from this NoE, since four partners (CERTH, CWI, Eurecom and UEP) already collaborated here. LinkedTV has a wider goal than just the focus on media analysis. K-Space did not have specific use cases while LinkedTV will further develop analysis and annotation technologies tailored for the needs of Connected Media Experiences and evaluated in three scenarios. |

| | |
|---|---|
| **Muscle** (FP6-NoE) — MUSCLE aims at establishing and fostering closer collaboration between research groups in multimedia data-mining and machine learning. The Network integrates the expertise of over 40 research groups working on image and video processing, speech and text analysis, statistics and machine learning. The goal is to explore the full potential of statistical learning and cross-modal interaction for the (semi-)automatic generation of robust metadata with high semantic value for multimedia documents. The project has a broad vision on democratic access to information and knowledge for all European citizens and it is quite focused on the full potential of machine learning and cross-modal interaction for the (semi-) automatic generation of metadata. http://www.muscle-noe.org | MUSCLE focused mainly on cross modal integration using offline machine learning techniques. LinkedTV will tackle cross-media integration and interaction using near real-time methods that do not lie on large scale offline learning. |
| **X-Media** (FP6-IP) — X-Media addresses the issue of knowledge management in complex distributed environments. It studies, develops and implements large scale methodologies and techniques for knowledge management able to support sharing and reuse of knowledge that is distributed in different media (images, documents and data) and repositories (data bases, knowledge bases, document repositories, etc.). http://www.x-media-project.org | In comparison to LinkedTV it employs a quite different application domain (automotive industry) and does not address further challenges of open Web-based search and retrieval of media. |
| **aceMedia** (FP6-IP) — The main technological objectives of aceMedia were to discover and exploit knowledge inherent to the content in order to make content more relevant to the user; to automate annotation at all levels; and to add functionality to ease content creation, transmission, search, access, consumption and re-use. In addition, available user and terminal profiles, the extracted semantic content descriptions and advanced mining methods were used to provide user and network adaptive transmission and terminal optimised rendering. http://www.acemedia.org/aceMedia | LinkedTV will build on the main outcome of this project, the aceMedia toolbox, while extending the technologies to fit with the needs of Connected Media Experiences and three specific scenarios. In particular, LinkedTV focuses on finer grained annotations of media fragments with individual concepts and on the integration of Web-based related media. |
| **MESH** (FP6-IP) — MESH is an Integrated Project whose main objective is to extract, compare and combine content from multiple multimedia news sources, automatically create advanced personalised multimedia summaries, syndicate summaries and content | MESH is focused on a single application domain (i.e. news) and on the spatial segmentation of news media, while LinkedTV develops a platform for generic multimedia services, implements three different scenarios and requires a finer grained spatio-temporal |

| | |
|---|---|
| based on the extracted semantic information, and provide end users with a "multimedia mesh" news navigation system. The goal of MESH is to develop an innovative platform for rapid and effective access & delivery of news. The MESH project was initiated with the vision to integrate semantic technologies into a setting that will bring the world of news closer to knowledge-enabled services. http://www.mesh-ip.eu/?Page=Project | segmentation of different types of media. |
| INTERMEDIA (FP6, NE) — Interactive Media with Personal Networked Devices. In this project the user is the point where multimedia applications and the means for interacting with them converge. Users are provided with a personalised interface and content independently of the particular set of physical devices they have available for interaction, and independently of the physical space. | The project focuses on the whole range of personal adaptations for multi-media sources onto different types of wearable devices for moving users. LinkedTV is not concerned with the needs of moving users and ad hoc connections. It applies personalization not in a wide range, but only when filtering concepts and content according to the individual profile. |
| **Live (FP6, CP)** — Staging of Media Events. The LIVE project seeks to change the linear approach to live sports broadcasting by delivering digital technologies and interactive content – through personalised on-demand sub-channels. | In LIVE the focus was solely on a live TV production of an inter-linked multi-channel TV bouquet. The channels were linked by hand of professional TV producers according to the flow of the sport events. For the manual (online) and semi-automatic (offline) annotation of broadcast content (live streams / archived clips) two dedicated tools were developed, which might be re-used and adapted for usage in LinkedTV. |

| | |
|---|---|
| **MediaCampaign** — Discovering, inter-relating and navigating cross-media campaign knowledge. MediaCampaign's scope was on discovering, inter-relating and navigating cross-media campaign knowledge and to automate a large degree of the detection and tracking of media campaigns on television, Internet and in the press. For the pilot system developed within the project the project focus was on a concrete example for a media campaign: advertisement campaigns. The project goals included the design and implementation of a specific media campaign ontology, cross relation of specific campaigns, algorithms for detecting advertisements over different media, audio analysis algorithms and the detection plus tracking of new campaigns. With regard to audio analysis work concentrated on segmentation, word spotting, dedicated speech to text modules and jingle recognition. | MediaCampaign relates to LinkedTV in the sense that both projects aim at the automation of the labeling process and interrelating media from multiple sources. LinkedTV will deploy the speech recognition toolkit developed partly in this project that is currently being used within Sound and Vision for audio analysis. LinkedTV differs from MediaCampaign mainly with respect to application domain (very specific user group in MediaCampaign), the analysis techniques addressed and the level of interlinking media sources: on the document level in MediaCampaign, on the concept/fragment level in LinkedTV. |
| **NM2** — New Media for a New Millennium created a variety of new media genres using broadband communication and interactive terminals. The project created new production tools for the media industry that allow the easy production of non-linear broadband media that can be personalised to suit the preferences of the individual user. Viewers are able to interact directly with the medium and influence what they see and hear according to their personal tastes and wishes. http://www.ist-nm2.org/ | The emphasis within NM2 was on the design of new genres and providing support for those creating the production. While users are able to select material according to their own wishes, all material is pre-determined by the content owner. LinkedTV will go beyond the boundaries of the published content and link to networked media available from other sources. All interaction points in NM2 are pre-specified, whereas in LinkedTV, active areas will be identified at view time, dynamically linked with other networked sources and a small set of relevant choices presented to the user to select from. |
| **porTiVity** — Rich Media Interactive TV services for portable and mobile devices. Developed a converged Rich Media iTV system which integrates broadcast and mobile broadband delivery to portables and mobiles aiming to enable the end-user to act on moving objects within TV programmes. porTiVity combined and further developed the technical achievements of past projects towards handheld TV. It offered tools and expertise to the broadcasters, playout-specialists, network operators as well as handheld terminal manufacturers. | In porTIVity, there are also selectable objects in AV material, however the identification of objects as well as their linkage to other content is done manually and in advance at the broadcaster side while LinkedTV focuses on a semi-automatic approach using intelligent approaches and re-using the Web metadata layer (Linked Data) as well as extending it with a Connected Media Layer. to provide an enhanced experience to the user by continuously retrieving Web content of interest related to the currently viewed material. Furthermore the project will offer more efficient and fast tools to the broadcaster in order to speed up the process of annotation. |
| **PrestoPRIME** develops practical solutions | Related to LinkedTV, PrestoPRIME will |

| | |
|---|---|
| for the long-term preservation of digital media objects, and find ways to increase access by integrating the media archive with European on-line digital libraries in a digital preservation framework.<br>The project aims to establish metadata interoperability between audiovisual archives, cultural heritage institutions, the Semantic Web and content portals, with services for metadata conversion and deployment. | create a metadata conversion and deployment toolkit, with a novel and efficient process for metadata vocabulary alignment. PrestoPRIME focuses on aligning manually created vocabularies that in many cases is limited to programme level descriptions. The Data Mining activities of LinkedTV (see 1.2.2) will go much further than processing these manually added labels. In connection with semantic concepts and entities extracted from the low-level audio-visual analysis, information from various outside sources will be collected and filtered. The user will be presented information relevant to the current spatio-temporal segment of a frame, which represents a much finer granularity. |
| **RUSHES** — Retrieval of multimedia semantic units for enhanced reusability: The overall aim of the RUSHES project is to design, implement, and validate a system for indexing, accessing and delivering raw, unedited audio-visual footage known in broadcasting industry as "rushes". The goal is to promote the reuse of such material, and especially its content in the production of new multimedia assets by offering semantic media search capabilities. | RUSHES aims at indexing and retrieval of content for a specific industrial application ("rushes" content management). As opposed to LinkedTV, RUSHES does not tackle the lack of integration between the television broadcast and the Web content, nor does it support the establishment of links between different yet related multimedia content items. The networked aspects of RUSHES are limited to the development of a collective annotation module, supporting the semi-automatic enrichment of tag-based media annotations, as opposed to the highly-developed Web dimension of LinkedTV, which will support the linking between networked media and individual objects in the video stream. |
| **Salero** — Semantic Audiovisual Entertainment Reusable Objects: aims at making cross media-production for games, movies and broadcast faster, better and cheaper by combining computer graphics, language technology, semantic web technologies as well as content based search and retrieval. | Salero also considers the creation of what they call « intelligent content » for broadcast media, however this is complex content created manually by experts rather than dynamically produced by tools re-using existing Web-based media. Since this content is annotated, it may be re-usable as part of the multimedia presentations generated by LinkedTV services. |
| **SEMEDIA (FP6, CP)** — Search Environments for Media. SEMEDIA aims to create new methods, environments and widely usable tools for media labelling, searching and retrieval from very large collections of heterogeneous data, building on and extending research in media technologies, web semantics, AI, Content Based IR and interface design. | Semedia also addresses search in streaming media but focuses on the support of the production process by finding similar scenes or specific light conditions and the automatic generation of an overview from a video.<br>In addition, LinkedTV will start much later and can therefore uptake the most recent research results regarding video analysis and similarity-based retrieval. |
| **CHORUS & CHORUS+** — CHORUS was | Partners of CHORUS and CHORUS+ |

| | |
|---|---|
| a Coordination Action which aimed at creating the conditions of mutual information and cross fertilisation between the projects that run under Strategic objective 2.6.3 (Advanced search technologies for digital audio-visual content) and beyond the IST initiative. CHORUS+ is the continuation of CHORUS in FP7. http://www.ist-chorus.org/ | (CERTH) participate in LinkedTV and therefore the project will benefit from the useful experience and the results of CHORUS / CHORUS+. More specifically, LinkedTV will reuse the state of the art reports on audio-visual search technologies produced by these coordination actions and also the user, market and gap analysis studies in order to prepare its exploitation and dissemination plans in the search sector. In addition, links with the CHORUS / CHORUS+ project clusters will be set up when necessary in order to collaborate and re-use existing technology. |
| **VidiVideo** — Improving the accessibility of video. http://www.vidi-video.it/ The VidiVideo project focused on the development of a semantic search engine for video. The main technology behind it was the development of a 1000 element thesaurus for automatically detecting semantic concepts in the audio-visual content. | LinkedTV will exploit the technological advances made within VidiVideo to semantic concept detection in video. CERTH was a partner of the VidiVideo project, where it lead the "Video Processing" Work Package. |
| **CASAM (FP7, CP)** — Computer-Aided Semantic Annotation of Multimedia. CASAM develops a solution for the double problem of expensive manual work and unexploited multimedia resources by automatical analysis of multimedia sources and provide systematic sets of descriptions for their content. | LinkedTV can use an ontology for the domain "Environment" to be developed in CASAM. CASAM clearly focuses on aspects relevant for the annotation of multimedia material in a specific domain, while LinkedTV will go beyond this process providing semi-automated indexing and content-based linking procedures to external sources. |
| **iMP** — intelligent metadata-driven processing and distribution of audiovisual media: will create architecture, workflow and applications for intelligent metadata-driven processing and distribution of digital movies and entertainment. | iMP aims to help media professionals better organize and process large scale digital media repositories. Hence its focus on the use of annotation is different than LinkedTV's goal of content aggregation and presentation. While there may be some co-operation possibilities in metadata-driven processing of digital media, LinkedTV considers more fine grained re-use of material in association with other content. |
| **INEM4U** — Interactive networked experiences in multimedia for you. The aim of iNEM4U is to facilitate enhanced multimedia experiences for individuals and communities. It will provide a rich and intuitive way for people to consume, share, interact with and communicate about multimedia content. It will flexibly combine media elements and value-added services from different worlds, such as broadcasting, IPTV, mobile and the Web, into one interactive multimedia experience. One aim | INEM4U aims at providing a seamless experience for the end user in terms of interacting with different devices. A major objective is to develop the infrastructure to allow content from different network and service environments to be accessible from mobile, IPTV, internet, and broadcast networks. This overlaps with a number of the objectives within LinkedTV, where the intention is to provide multi-device access to content. Within LinkedTV more emphasis is given to identifying connection points within |

| | |
|---|---|
| is to increase the feeling of "connectedness" with other people or the perception of "involvement" in a particular shared experience. http://www.inem4u.eu/ | media and allowing these to draw in related media assets from multiple networked media sources. |
| **INSEMTIVES** — Incentives for semantics: bridge the gap between human and computational intelligence in the current semantic content authoring R&D landscape | Insemtives focuses on supporting the generation of semantic annotations through incentive-based tools. This project has just begun (April 1, 2009) and may complement LinkedTV through sharing research both in the broadcaster annotation tool and end user interfaces. |
| **MYMEDIA (FP7,CP)** — Dynamic personalisation of Multimedia. The resulting system should allow easy integration of multiple content catalogues and recommender algorithms in a single system and provide technology for ranking the content based on personal preferences. | MYMEDIA emphasizes on the personalization of content consumption and recommendation of relevant content to the user. LinkedTV can benefit from the user models developed within MYMEDIA and the field trials for the project context-aware personalization research activities and functionalities. However, MYMEDIA does not fully exploit content-based analysis techniques for annotation and does not provide automated, personalized linking with relevant content from the Web. |
| **PetaMedia (FP7, NOE)** — P2P Tagged Media. The project is directed towards the synergetic combination of user-based collaborative tagging, peer-to-peer networks and multimedia content analysis, and towards identification and exploration of potentials and limitations of combined tagging, MCA and SP2P concepts. | The Call 2 project PetaMedia shares with the LinkedTV proposal the effort in exploiting complementary textual and social-network data for indexing multimedia. However, while PetaMedia primarily deals with the back-end algorithms for media indexing, LinkedTV pays significant attention to the user-facing front end. Furthermore, PetaMedia as an NoE aims at integrating existing basic research at different workplaces, while LinkedTV, as an IP, will lead to a functional software architecture, joint elaborated use cases and business models required for its exploitation. The projects are thus complementary at instrument level. |
| **NoTube** — Networks and Ontologies for the Transformation and Unification of Broadcasting and the Internet | While NoTube focuses on the personalisation of TV broadcasts at an atomic level (the program itself or temporal segments thereof), we will focus at a much lower granularity (individual spatio-temporal segments). Likewise, the association to related content is at an atomic level in NoTube, as opposed to LinkedTV's aim to display related media segments dynamically and in an intuitive fashion to the viewer. |
| **MultimediaN** — Dutch national project. http://www.multimedian.nl/en/home.php. Research issues tackled within MultimediaN | Many of the results from MultimediaN are relevant to LinkedTV. In particular, LinkedTV will build directly on results on user |

| | |
|---|---|
| were addressed in a number of sub-projects: Learning Features, Multimodal Interaction, Ambient Multimedia Databases, Semantic Multimedia Access, Professional's Dashboard, Video At Your Fingertips, E-Culture, and PERsonal Information Services. These covered a wide range of results including the detection of regions within images and video, and associating these with semantic concepts. User issues for providing access to media resources were also addressed. | interaction with semantically annotated cultural heritage repositories in the E-Culture project. Here information exploration tasks were investigated, in particular how image assets relate in different ways to concepts that are of interest to the user. LinkedTV will go beyond these tasks to investigate how users can navigate through linked video assets. LinkedTV will also improve on the video region detection techniques to allow real-time tracing of moving regions. |
| **Quaero** — French national programme. Quaero is a collaborative research and development program, centered at developing multimedia and multilingual indexing and management tools for professional and general public applications such as the automatic analysis, classification, extraction and exploitation of information. The research aims to facilitate the extraction of information in unlimited quantities of multimedia and multilingual documents, including written texts, speech and music audio files, and images and videos. Quaero was created to respond to new needs for the general public and professional use, and new challenges in multimedia content analysis resulting from the explosion of various information types and sources in digital form, available to everyone via personal computers, television and handheld terminals. http://www.quaero.org/modules/movie/scenes/home | Both projects tackle multimedia indexing. LinkedTV is more oriented toward the end user (content personalization) with the main goal of developing new experiences of IP-based TV and video viewing, enabling the browsing of objects within TV and video programming and the dynamic and personalized on-screen integration of Web-based media related to the represented concepts. |