Connected Media Experiences: ConnectME

Proposal full title:
Connected Media Experiences

Proposal acronym:
ConnectME

Type of funding scheme:
Integrated Project (IP)

Work programme topics addressed:

Challenge 4: Objective ICT-2009.4.3: Intelligent Information Management
Target outcome: a) Capturing tractable information and b) Delivering pertinent information

Name of the coordinating person: Dr Rüdiger Klein, Fraunhofer IAIS
e-mail: ruediger.klein@iais.fraunhofer.de
fax: +49 2241 14 2342

List of participants:

Partici-pant 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	University of Mons	UMONS	Belgium	University
5	Centrum Wiskunde & Informatica	CWI	Holland	Public Research Center
6	Eurecom	EURECOM	France	Public Research Center
7	University of St Gallen	UNISG	Switzerland	University
8	Semantic Technology Institute International	STI	Austria	Public Research Center
9	Condat	CONDAT	Germany	SME
10	Sound and Vision	S+V	Holland	Public Company
11	Deutsche Welle	DW	Germany	Public Broadcaster
12	Thomson Grass Valley France	TGV	France	Industry

Abstract

ConnectME is an integrated practical approach
towards capturing and delivering the wealth of information in the rich multimedia Web

ConnectME stands for Connected Media Experiences. In the next years multi media content, especially video based material will play a much more central role on the Web than today. The amount of multimedia material distributed via IP and the Web, such as television broadcasts or user-generated content will scale up to unimaginable levels. Currently, this material is much less usable than text based information. Media is tagged in social websites and can be found with traditional search engines. But neither approach provides sufficient access to the rich information content of multi media material, and none of these approaches can deliver the necessary precision, granularity, and personalization to make that information truly accessible in a useful manner. **The core problem is that multi media material is currently isolated: it is not connected across providers and services, and it is not adequately interlinked with other information on the Web.**

In addition to current providers, ever more companies, organizations, and social groups will provide multi media information – much in the same way as text based information is provided today. Companies will publish user hand-"books" about their technical products, teachers their lessons, artists groups their performances in multi media form. They all make use of the human capability to "process" multi media material much more easily and intuitively than text based information. However, the value of these huge scales of online media is limited by that media's isolation on the World Wide Web.

ConnectME will enable providers and end users to benefit from the future scale and ubiquity of multi media material in an intelligent and efficient way. ConnectME does not have to start from scratch – many pieces necessary in this context do exist today. What is still missing is an overall integrational approach bringing the pieces together in a coherent way:

- Multi media material will be annotated in a fine grained and precise way using provider information, Web based background knowledge, social networks, and user centric personal information.
- It will be interlinked in this way with multi media information and other Web based information.
- Users will be enabled to retrieve Web based multi media information intuitively based on what they see and hear in audiovisual streams.

ConnectME will develop the unifying conceptual framework, design and implement a prototypical platform, and validate the overall approach with three typical use cases. The infrastructure behind ConnectME will be an end-to-end platform to host those added-value services over different networks, providing the common required functionality of each service. This is based on finely weaving media content together on the basis of generated metadata at the individual object level.

To achieve this, several research challenges must be overcome. Multimedia analysis tools and methods must scale to meet future data size, and annotation schema must be able to capture richly the relationship between media and what it represents while being amenable to scalable reasoning. New services must be built able to sort through the huge amounts of generated annotation to select, adapt and package related media into informational presentations for the user. Techniques such as computational attention and behavioural tracking are needed to innovatively enable systems to select the concepts and content most useful and of interest to individual viewers, based on what they watch on screen and how they interact with the content. Finally, user interfaces must be provided so that the selection of on-screen objects and the browsing of the associated content can be done in a non-disruptive and intuitive fashion.

Executive Summary

ConnectME will allow us to **seamlessly connect multimedia information** on the Web by integrating networked media analysis, personalisation and presentation technologies within an integrated and coherent framework. Rich multi media information will be much more usable in this way. ConnectME will develop a comprehensive methodology, as the basis for the ConnectME 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 ConnectME platform will support the whole Intelligent Information Management process from end to end in an integrated and coherent way. The result will make Web-based multi media information more useful and valuable, and it will open completely new areas of application for accessing information on the Web via rich media experiences.

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. ConnectME will build on existing research projects like LarKC for efficient large scale processing capabilities on the Web, OKKAM for providing unique entity annotation techniques, NoTube for its multi media processing approach, and integrate them in a coherent and holistic approach for future media and the Web.

Through business modelling and industrial exploitation partners, Connected Media Experiences will become a core technology for the Web of the future, forming the basis for advanced information multimedia applications. ConnectME will provide the technologies needed in the broadcast and telecommunication industries to enable innovative services which can create new market opportunities in Europe for SMEs and protect and improve the competitiveness of large traditional market players in an increasingly fractured future global media landscape. The ConnectME approach will also be adaptable to other contexts such as mobile, 3D Internet or virtual worlds. ConnectME demonstrators will be used to present Connected Media Experiences 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.

Section 1. Project Description

Concept and objectives

The World Wide Web already today has completely changed the way we are providing and using information. In the last years 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 uploading and viewing on Web based portals like MySpace or YouTube is growing in usage at a fast rate (YouTube recieves 20 hours of new video every minute and serves one billion videos each day[1]), spreading out onto multiple devices, while media content owners such as cultural archives and TV broadcasters shift to the Web as distribution channel for their material. But the information in this rapidly expanding rich media Web is not as accessible and usable today as text based information.

The reason is that there are still important gaps and breaks in the current rich media Web. Audio-visual content is not interlinked with other material on the Web. Content links are produced manually, with a resulting lack of scale and precision. Web-video is video delivered over the Web, but barely making use of the potential and opportunities possible through the huge scale of related information which is available over the same channel. Web-TV is not much different from the digital TV piped to the user by cable or satellite - though some first examples seek to apply new features based on Web interactivity (see section 1.2). The vision of Connected Media Experiences 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 or even more will be used tomorrow: interlinked with each other, with other kinds of information, intelligently searchable, personalized, and accessible everywhere and every time.

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.

Connected Media is a central element of the vision for the future rich media Web. It will allow us to use audio-visual information in the same interlinked way as text can be used on the current Web. Connected Media will 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 to read boring manuals if a short video sequence allows people to find out what they need? Why to 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).

There is another important difference between text based information and videos: the content of text is explicit whereas video content needs perception and interpretation. Of course, also

[1] http://www.wired.com/epicenter/2009/10/youtube-over-one-billion-videos-served-per-day/ Oct. 2009

text based content needs interpretation and embedding in a context by users. In video based information this process is much more complex and related to the user. The amount of information in a single video frame can be quite complex, and needs filtering and personalisation to the user's interests, which need to be captured in an implicit and intuitive manner. For this, attention tracking techniques can help systems learn which objects in media are most relevant to a given user.

HyperVideo 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. ConnectME will pursue a *practical* approach towards Connected Media which is now manageable.

ConnectME will aim at two main and related issues: interlinking and presentation. We will provide interlinking of video based information with all other kinds of Web based material, and a completely new user-centric presentation and interaction approach. A fine-grained annotation of video content can be achieved by 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. ConnectME 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.

Another important source of information to be used in fine grained video annotations is user feedback. It comes from two channels: social groups and user experiments. Collecting information about user group behaviour enables us to find out how video content is used (and interpreted by users). Data mining techniques will enable us to use this information for user centric video content interpretations. Additionally, psychological experiments on user behaviour and perception will provide more detailed information about video content perception and interpretation by different groups of users.

ConnectME will concentrate on three tightly interconnected main goals towards a practical approach to a future of Connected Media:

- **ConnectME will develop a comprehensive methodology for a rich media Web with large scale information capturing, reasoning and management** to be used in the course of fine grained video annotation for the Web. This provides the basis for interlinking, search, personalization, and contextualization.

- **We will create the ConnectME platform enabling end-to-end support of large scale rich media on the Web**. It will 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.

- **Two 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.

ConnectME will address the challenge of handling media analysis, annotation, selection and packaging into multimedia presentations at the scale of the Web. To summarize, here are the

problem areas addressed by ConnectME, the innovations of this project, the output foreseen in the work programme and the benefit to users:

Table 1: ConnectME innovation

Problem Area	Innovation	Output	User benefits
Current audiovisual analysis approaches do not provide fine grained enough identification of individual concepts in the material	Novel semi- or fully-automatic analysis algorithms to assist in 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	Tool for concept-level annotation of fragments of audiovisual material (WP1), incorporating conceptual annotation support through complementary Web data mining (WP2)	One can more efficiently and effectively annotate media content.
Current platforms are not designed to analyse and integrate large scale multi-media materials including video	Planned deployment of the LARKc platform for storage and retrieval of large amounts of meta data. LARKc is developed in the project www.larkc.eu which has already provided promising results.	Connected media platform using the underlying LARKc systems to store and retrieve data. Alternatively one of the RDF triplet storages such as the Linda system.[2]	Increased performance during navigation and search through large data storages.
Current multimedia annotation schemes are not sufficient for unambiguous retrieval and packaging of associated Web content	New annotation schemes for Connected 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	One can share and re-use annotations
Current multimedia-based information services over the Web 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	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	Providers of added value services can offer a wider range of new innovative services at lower cost (re-using Web content) and with more personalization (increasing potential commercialization).

[2] www.w3.org/2006/02/DERI.pdf

	be generated on the fly	ns to user activity, behaviour and context (WP4)	
Current access paradigms to audiovisual material, including online video or TV, do not present a tight integration with associated content in the global Web	New services will be enabled where consumers of audiovisual material may have instant access to information about concepts represented in the viewed material	A ConnectME platform will integrate all components produced in the project (WP5) and provide the basis for the implementation of individual scenarios (WP6)	Content owners and providers 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 ConnectME 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 converted into enterpreneurship and new market opportunities	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 plans and promoting exploitation of project results (WP8).	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.

1.1.1 Project description

ConnectME stands for Connected Media Experiences. This refers to services which provide information to users through the dynamic interlinking of associated media, enhancing the video experience.

In the future, the amount of multimedia material distributed via IP and the Web, such as television broadcasts or user-generated content, will scale up to unimaginable levels. This material is, as yet, not connected across providers and services, even though many added-value and commercializable services could be built on mash-ups of media at a granular level, e.g.

- IPTV services providing digital TV viewers intuitive and dynamic access to Web-based information about objects in the broadcast stream;
- Web services linking uploaded video material to related Web content and TV programming;
- Archives, such as those in the cultural heritage domain, benefit from deep linking with the wider media archive of the Web.

In addition to current providers, in the future many companies, organizations, and social groups will provide multi media information – much in the same way as text based information is provided today. Companies will publish user hand-“books” about their technical products, teachers their lessons, artists groups their performances in multi media form. They all make use of the human capability to “process” multi media material much easier then just text based information.

The infrastructure behind ConnectME will be an end-to-end platform to host those added-value services over different networks, providing the common required functionality of each service (Figure 1 below).

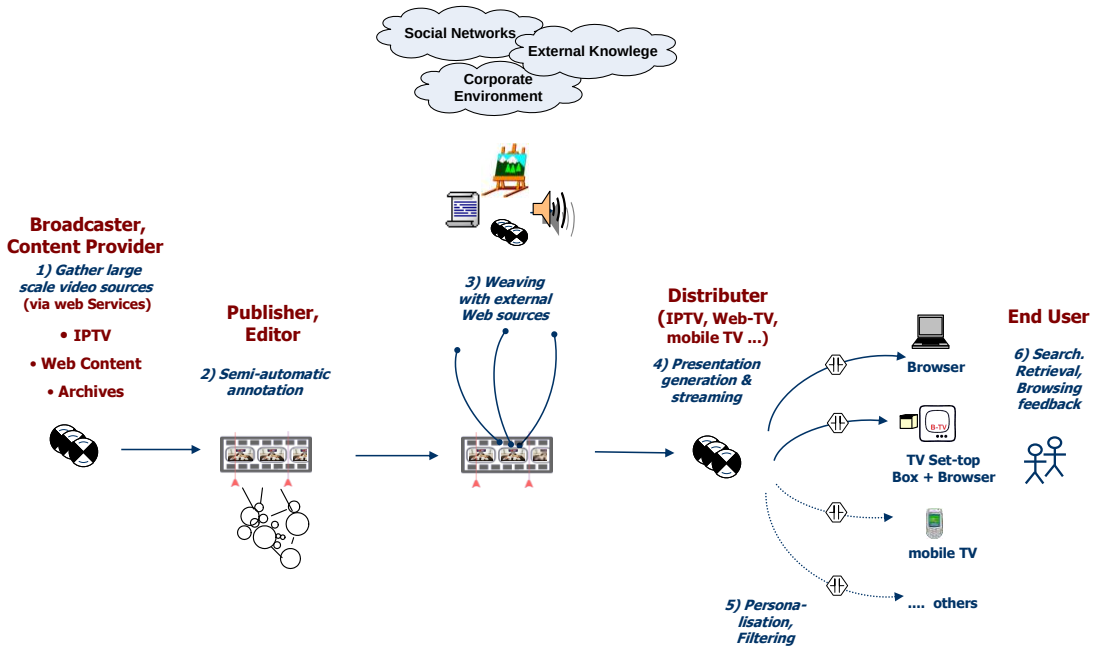


Figure 1: The core steps of creating Connected Media Experiences

This is based on finely weaving media content together on the basis of generated metadata at the individual object level and in different shells of context (Figure 2 below): data in the local environment, external knowledge on the Web, and user-generated content on the Social Web (otherwise called Web 2.0).

Basic Concept

Audio-visual information can be attractive and effective in many cases: from typical industrial and business 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 information in audio-visual material in an effective and efficient way – adapted to the many different kinds of usage of audio-visual information.

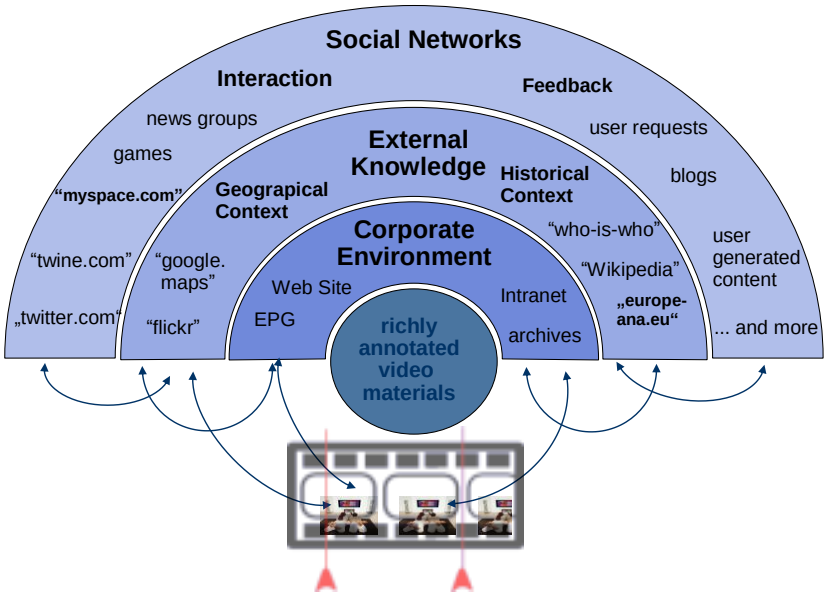


Figure 2: Weaving Connected Media into a Connected Media Experience

Viewers may want to find out the following:

Figure 3: Sample ConnectME-enhanced audio-video information

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

- Fine-grained annotation of audio-visual information at the scale of the Web 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 ConnectME 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. 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 their way around the complex, large scale, networked information extracted from the Web. 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.

1.1.2 Scenarios

We will demonstrate our methodology of weaving of connected information across the Web into Connected Media Experiences and the capabilities of our ConnectME platform through two scenarios, both representing different aspects of the value and potential of the rich media Web extended by a Linked Media layer which allows the interlinking of media relating to same and related concepts (see also chapter 1.2.1 for a more detailed description):

A **business** scenario will showcase how Connected Media Experiences can empower knowledge workers in collaborative business and industrial settings through instantaneous information access, whether it is business people making important decisions or customers seeking to find the right solution to their problems.

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.

Both scenarios provide quite different requirements to the ConnectME 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 business scenario will frequently need precise and fine grained annotations. There is a lot of information available for this purpose in business and industrial environments which provide the needed quality and granularity. The ConnectME platform can make use of all information available in the course of media production including the general business and technical information related to the case at hand. The search functions will need high precision. The ConnectME user interface has to be tightly integrated into the general business user interface.

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 ConnectME

user interface should support the user to manage the rich information associated to cultural content.

1.1.3 Future Web as Connected Media

ConnectME will make possible *connected media* through a Web-wide metadata layer, which links media to concepts, and links concepts together. The hypertext Web evolves into a hypervideo Connected Media Platform, where the access paradigm is not the browser but an audio-visual programming channel from which the viewer can access concepts and browse content.

The ConnectME platform can enable this new paradigm for interactive video (*hypervideo*) with dynamic and personalized content enrichment on any online media device, hence making Connected Media Experiences the ubiquitous and intuitive information delivery means of choice in the future digital society.

1.2 Project scenarios

To illustrate the potential of Connected Media Experiences, we will build demonstrators around two scenarios which will represent different requirements to our methodology and tools, different target audiences, content topics and access contexts.

These scenarios arise out of the recognition that audio-visual information is not an isolated medium on the Web, 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. Content trends are leading them to become cross-media production companies. Audio-visual content will completely be integrated into the 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 owned by consortium members Sound and Vision and Deutsche Welle) 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 Semantic Web technologies (Open Linked Data) which can enable concept-based search and browsing for information.

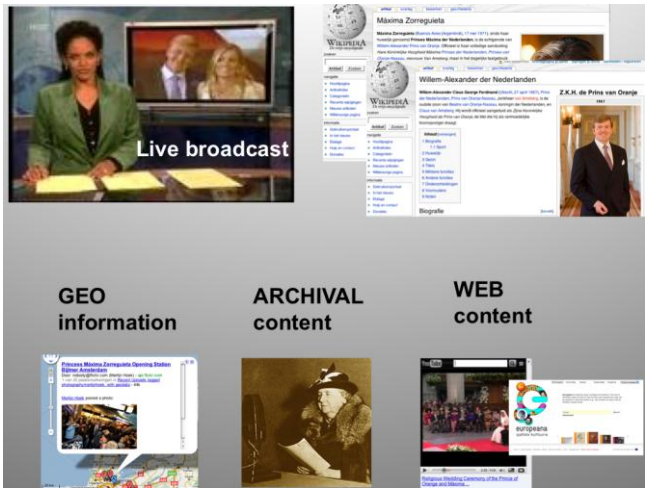


Figure 4: Sample Components of a Connected Media Experience

1.2.1 Scenario 1: Using enhanced audio-video information in business environments

Andy is an investor interested in companies with promising economic prospects. There is no doubt that renewable energy already is and will be a booming market within the next decade. Andy wants to collect preliminary information about this sector before he considers pursuing an investment idea further. Sitting in his office in front of his laptop he is watching a program on alternative energy sources. One video sequence is about wind power generation on the shores of Northern Germany. The companies behind the wind turbines are of particular interest to him. He would like to see a short history and profile of the company (text-based), promotional materials (images/video) as well as recent news material that mentioned the company (text/audio/video) which should be presented in a chronological manner. Andy furthermore wants to find out about the turnover of the companies involved in order to decide whether they are worth an investment. The political situation and context regarding environmental issues have a great impact on his decisions, too (e.g. the situation concerning subsidies on the regional, national and European level, or the regulatory framework). Hence, Andy wants to obtain all this information as quickly and efficiently as possible in order to be able to make an informed, facts-based decision.

In addition to the above, Andy also wants to evaluate the popularity of ecological measures by analyzing user-generated content surrounding the service in question. By amalgamating what kind of ecological energy production appeals to the general public the most, he is in a better position to advise third parties about possible sponsoring opportunities because of its standing with public opinion. 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. ConnectME 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 ConnectME 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 offerings of his competitors.

Laura is an engineering student who wants to learn more about green energy in general and ways in which she can help consumers to do more to protect the environment. At her PC she is watching a program on wind turbines. She wants to understand (on a basic level) how they function, how they compare to other forms of energy production (advantages/disadvantages), and where they are already deployed (possibly to have a look at them). She furthermore wants to find out which electricity companies in a special area already offer energy based on wind power, and compare prices. 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 like she is, and do so both on a regional as well as national level. The tools and services to be developed in ConnectME, 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 go with her ecological conscience. Through ConnectME, Laura is in a position to make more informed decisions.

This business scenario is characterized by the need of precise and fine grained annotations. For the content producers there is a lot of information available for this purpose in business and industrial environments which provide the needed quality and granularity. The ConnectME 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 Andy, John, and Laura the search functions will need high precision. The ConnectME user interface will be tightly integrated into the general business user interface.

Concept	**Flender A106B**	**Ulrich Bez**
Class	Wind Turbine	Person
Related content	**For an engineer (browse on the concept):** Turbine specifications Diagrams Technical manuals **For a member of the public (browse on the class):** General description „what is a wind turbine?“ Diagrams „how does it work?“ Images/videos of turbines	**For a business person:** Ulrich's position in the company His public CV LinkedIn/XING profile **For a member of the public:** Other texts, photos, videos with Ulrich Bez

Browsing facets	**To the company** Flender AG **To other types of wind turbine** **To other forms of energy production**	**To the company** Flender AG **To other employees at Flender AG**

Figure 5: Example for the business scenario

For this scenario, we have access to the video archives of Deutsche Welle, Germany's national broadcaster. The programmes Made in Germany – the Business Magazine and Tomorrow Today – the Science Magazine will be used in particular for this scenario.

1.2.2 Scenario 2: Using enhanced audio-visual information in cultural heritage learning

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 ConnectME 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 by Deutsche Welle 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, ConnectME 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, ConnectME 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 ConnectME 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. ConnectME 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 DW 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.

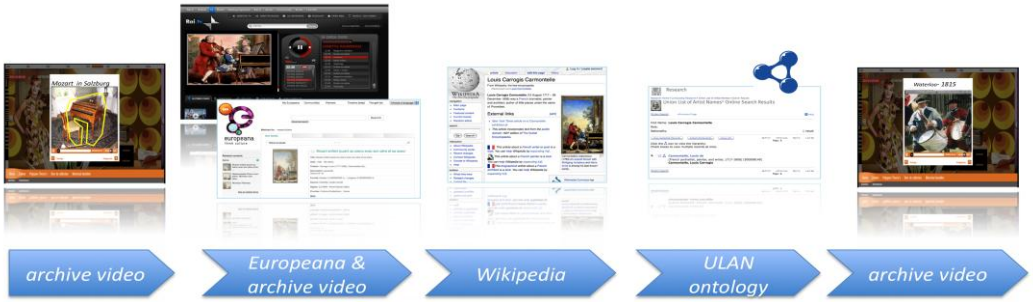


Figure 6: Content used in the documentary scenario

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

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

For this scenario, we have access to the cultural heritage archive of Sound and Vision, which hosts 70% of the Dutch national audio-visual heritage. In total, with around 700,000 hours of television, radio, music and film; making Sound and Vision one of the largest audiovisual archives in Europe.

1.3 Project objectives

The Connected Media Experiences IP proposal aims at an innovative, integrative, and practical approach towards connecting media across the Web, which will be tackled via a set of tightly interrelated key scientific and technological research goals and measurable project objectives:

- The ConnectME 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 ConnectME services which can scale to the size of the Web.
- The ConnectME methodology and platform support the Connected 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) re-using the huge and growing amount of machine-processable descriptions of media objects and general concepts on the Web to enable a "Linked Media" metadata layer on the Web, complementary to the Open Linked Data activity.
- 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 in 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 different kinds of users and scenarios. This involves developing user interfaces such that the selection of on-screen objects and the browsing of the associated content can be done in a non-disruptive and intuitive fashion.
- We will support users to manage the huge amount of interlinked multi-media information in a way tailored to their requirements and needs. This will include exploring the use of computational attention and behavioural tracking to innovatively enable systems to select the concepts and content most useful and of interest to individual viewers, based on what they watch on screen and how they interact with the content.
- The ConnectME 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 ConnectME 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 ConnectME methodology and our platform will be evaluated in our representative 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 industry participation, ConnectME will transfer its results to the content archiving, broadcasting and content delivery platform industries to enable new and innovative Web-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 rich media Web landscape.

1.3.1 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
Develop media analysis approaches for concept identification in video that are highly accurate and highly time-efficient, suitable for use with large-scale datasets.	1	(1) Automatic detection and tracking of objects or regions of interest (2) Automatic object and scene labelling, resulting in fine-grained semantic descriptions of scenes (3) Exploitation of compressed video information and currently unexploited hardware resources, to enable analysis in real-time (4) Exploitation of metadata for	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 and scalability of the developed technologies over a large-scale dataset

		linking of detected objects / regions to semantics / concepts	
Enable Web-scale hypervideo through the conceptual annotation of media	2	(1) A URI-based mechanism for addressing fragments of media content (2) Lightweight annotation schemes interoperable with most of the multimedia metadata standards (3) Web-scale data mining for refinement of multimedia processing (4) Acquisition of additional content from the Web	Client-side and server-side implementation of media fragments that conform to the Test Cases developed in the W3C Media Fragments WG 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 Ground-truth data setup, benchmark and evaluation campaign comparing existing information extraction tools and assessing the added value and scalability of novel disambiguation algorithms
Develop intuitive interfaces for interacting with Web-based media with respect to concept browsing and content presentation in video	3	(1) Understanding of user tasks for accessing linked information from audio-visual media. (2) Creation of interactive presentation engine for supporting tasks in a concept-based navigation environment. (3) Development of design guidelines for user interaction with linked information in a media network.	(1) Sufficient functional and user interface requirements to allow design of supporting software in the context of the project scenarios. (2) User environment satisfies identified requirements and can be incorporated into the overall ConnectME environment using scenario data. (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	4	(1) Unobtrusive detection of user's interests through their interaction activity and physical behaviour (2) User behaviour pattern recognition and knowledge pulling for	Manual or semi-automatic content and concept ratings, design of evaluation metrics and methodologies, user study

		contextual personalization (3) Innovative techniques on concept and content delivery adaptation according to computational attention / behaviour tracking	
Create an end-to-end platform for ConnectME services	5	(1) Platform deployable in heterogeneous environments (2) Flexible integration of all components (3) Seamless cooperation of all components to generate annotated and interlinked media presentations (4) Provide interactive user interface with personalization via user activity and behaviour tracking	(1) Evaluation of each component in the corresponding WPs 1-4 (2) Systematic verification of all features from end- to–end in WP5 (3) Evaluation from the user perspective by the trials in WP 6
Realize two scenarios where ConnectME services are used to enhance the Web-based TV or video experience	6	(1) A documentary scenario (2) An environmental scenario	Internal qualitative evaluation by the partners User trials using members of the target community Public feedback to scenario demonstrations
Disseminate project results in the community	7	(1) Conference papers (2) Journal articles (3) Workshops on ConnectME research topics (4) Standardisation actions at W3C etc.	Number of publications Global standing of the conferences and journals where ConnectME results are presented Attendance at and results of ConnectME workshops Contribution to standards
Exploit project results commercially	8	(1) Market surveys (2) Business plans (3) Exploitation plans	Market reports on potential for ConnectME services and technology Business plans taken up in commercial ventures Exploitation by ConnectME partners and others

1.3.2 Relation to call

ConnectME focuses on the target outcomes **a) Capturing tractable information** and **b) Delivering pertinent information** in the Strategic Objective 4.3 entitled Intelligent Information Management. As we will see, ConnectME results also have a relevance to further outcomes c) Collaboration and decision support and d) Personal sphere. Our focus on integrating and extending cutting edge research as well as channels to effectively exploit results commercially make ConnectME also a contribution to e) Impact and S&T leadership.

Target outcome a) seeks technologies to acquire, analyse and categorize extremely large, rapidly evolving and potentially conflicting and incomplete amounts of information. The ConnectME platform is built to sort through both seed audio-visual material from large archives/Web streams and associated media on the Web. It will be developed with scalability, dynamics and inconsistency of data in mind, using the latest techniques developed and research done within the consortium on aspects such as multimedia retrieval, Web mining, large scale reasoning and mediation between (possibly incomplete or inconsistent) media annotations. Based on an industry-grade Application Service Bus built by the leading electronics company Thomson Grass Valley, video and metadata handling infrastructure supplied by SME CONDAT who works with leading TV broadcasters in Europe and mature multimedia tools and techniques developed in previous EU projects and extended by ConnectME by our partners, the ConnectME platform will be *robust* and *performant*.

Technologies (that) *will extract, correlate and integrate data from diverse sources and formats* is the raison d'etre of ConnectME: to extract media from different sources which are dynamically interlinked at Web scale through conceptual representation and relationships and relate them meaningfully in spatial and temporal space as an information multimedia presentation.

Target outcome b) seeks usable and customisable systems to improve the efficiency of the information lifecycle. The ConnectME platform is aimed at server and client side solutions to information management, either in the selection and packaging of related multimedia resources from diverse sources (archives, streams, Web) for the information-providing service or in the browsing and presentation of those resources to the information-seeking user. Innovative interface design and personalisation (activity processing, attention tracking) techniques will make for a human-usable and customisable system. Maximised automation of media processing and annotation in the back end will aid its usability for broadcasters and other media owners, combined with customisation in terms of selected processing and annotation components, and media and Web sources.

Browsing concepts in the ConnectME player and selecting them for further information represents a *goal-dependent search* which triggers for the ConnectME platform back-end the *acquisition, structuring and aggegration of relevant local, remote and streaming resources*. Large scale reasoning of multimedia annotation data across the Web will be enabled by combining the framework developed in the EU project LarKC with the skills of the consortium in multimedia semantics. For the end user, this means the ability of *navigation, manipulation and consumption of digital information* in an intuitive and structured manner (in response to audio and visual cues in AV material), making use of innovative and cutting edge *adaptive user-information interactions based on ... human perception and attention* (eye tracking technology will provide attention data to an intelligent processing unit combining user profile and previous activity models to effectively filter concepts selected for browsing and the media selected for information provision).

ConnectME scenarios will demonstrate supporting business (environmental, in our case) and scientific (cultural heritage, in our case) communities in mining information referenced in AV material in a dynamic and instantaneous manner, helping them to make better use of

broadcast and archive material in their intelligence gathering and research; hence ConnectME has a relevance to target outcome c) Collaboration and Decision Support Tools.

ConnectME scenarios will also demonstrate supporting individuals in acquiring information during consumption of AV material in an intuitive and personalised manner, helping them *visualise and interpret* information found in audio and visual cues in AV material. The ConnectME platform underlies services for the *provision of personalised and context-dependent information from multiple sources*; hence ConnectME has a relevance to target outcome d) Personal Sphere.

Finally, it is a target outcome to ensure Impact in the professional community and S&T Leadership in Europe. ConnectME has a clear focus to achieve both: exploitation measures seek to ensure ConnectME results lead to real services based on multimedia information delivery to the European citizen, establishing European industry as a leader in this market. The existing excellence in multimedia and semantics research in Europe will be built upon and made actionable in ConnectME with concrete systems, tools and services, contributing to the further S&T leadership of Europe in this area.

Intelligent Information Management and ConnectME

Our concern in ConnectME is that media on the Web is not yet connected, neither to the information that it communicates nor to other media which is related. 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 desirable to be able to associate fragments of the consumed media to other, conceptually related, media on the network in order to build up a set of information around a particular subject in the media. However, traditional search and retrieval, such as entering a text string to find media, is not intuitive in the context of watching Web-based video or TV. Rather, ConnectME 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.

Hence ConnectME will develop a *platform for capturing and delivering the wealth of information in the rich media Web*: (1) for creation of ConnectME 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; (2) for supporting the ConnectME platform we will also specify guidelines and schemas for a Web-scale 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.

However, fundamentally ConnectME is about enabling a next generation of services which go beyond classical multimedia search and retrieval, and intuitively offer users **enhanced access to knowledge and presentation of information in a multimedial manner. We see this as the creation of a new hypervideo experience for browsing the rich media Web,** in a similar manner to how hypertext was the instigator of the success of the first generation of the World Wide Web.

1.4 Progress beyond the state-of-the-art

ConnectME targets the state of the art in highly scalable, domain independent multimedia analysis, annotation, personalisation and presentation. We aim to shift from a state of the art in which individual, specific multimedia services are hard coded for known domains, to an Internet of dynamic, scalable Web-based multimedia information services composed of interlinked, related media – what we refer to as Connected Media Experiences.

In our application to Web-based TV and video, we propose a system that will provide responsive, innovative content aggregation that goes beyond any offerings that are available at present. Quite like hypertext was the key to the success of the World Wide Web by its one-click link-following navigation paradigm, we argue, in the application of ConnectME functionality across networks and devices, that hypervideo will be realised with the potential to become key to the success of the rich media Web (for details on this position see also http://en.wikipedia.org/wiki/Hypervideo).

1.4.1 Web-based interactive TV/video

The model of interaction foreseen in ConnectME 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.

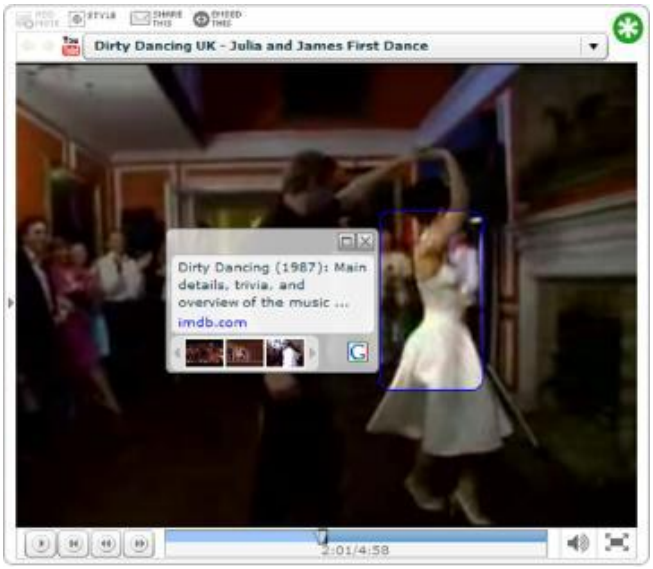


Figure 9: Asterbot allows the user to click on automatically detected salient regions and acquire information about related video clips or text in the web

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.

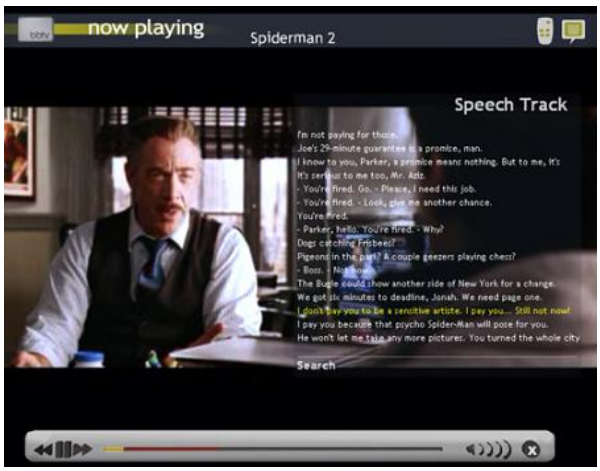


Figure 10: Blinkx BBTV interface

What ConnectME aims for is a more intuitive and automated approach that offer the video consumer dynamic and personalised access to associated content based on concepts in the video. The ConnectME 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.

1.4.2 IP-based Web and TV convergence

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.

The current trend in IPTV is towards Web integration 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.

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

Another trend is personalisation, with content recommendation and EPG personalisation making the TV experience more relevant to individual viewers. The Dutch project iFanzy[3], and now an EU project NoTube[4], research how to improve TV personalisation even further through using semantic metadata. These complement ConnectME as their focus is on complete TV programs, and the personalization takes place in the Electronic Programme Guide (EPG), while ConnectME looks more finely grained at personalization *within* the video stream at the level of individual video objects.

Finally, 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.

These trends do not 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

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

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.

ConnectME will hence help push Web-based TV delivery 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) interactive video player will provide the means for viewers to select objects “in-TV”; (2) the Connected Media infrastructure on the Web will provide the means to collect relevant media associated to the concept selected; (3) ConnectME 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 ConnectME 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.

1.4.3 Multimedia analysis

State-of-the-art technologies in the field of intelligent content processing and management developed in the scope of current research ([5], [6]) and recent European projects like aceMedia [1], X-Media [2], MUSCLE [3], MESH [4], 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 and cross-media mining applications are currently emerging as a result of these projects; 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 process. Furthermore the implemented fusion processes are mostly based on combining the single-media results rather than on continuous recursive cross-media interaction.

Large scale content processing 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 solely 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 [10][5] that incorporates image retrieval and face recognition for searching people and products on the web. Video sharing, as implemented by YouTube [11], put emphasis on the need for Content Based Image Retrieval (CBIR) on multimedia data (video, audio and text). The notion of *hypervideo*, where users are allowed to click on various aspects of a video to pull up additional information and links to external sites, is used to characterize tools like Asterpix [12], Videoclix [13], Hypersoap [15], Klickable [14], Overlay.TV [16] and Blinkx [17], 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.

[5] Other examples are: Google Similar Images (http://similar-images.googlelabs.com/), Bing Similar Images (http://www.bing.com), Xcavator (http://www.xcavator.net/) , Incogna (http://www.incogna.com), Tiltomo (http://www.tiltomo.com/) and also reverse image search engines, like TinEye (http://tineye.com/) or BYO (http://labs.ideeinc.com/upload/), where people may upload their own images as queries.

ConnectME will exploit state-of-the-art tools, similar to the core idea of Asterpix and the high quality viewing experience of Videoclix, and extend them in order to allow for minimal human intervention. Intelligent multimedia processing in the scope of the project will tackle the main technological barriers of how to *detect* regions of interest in a computationally efficient way with respect to both large scale datasets and timing, *track* them over time, identify and finally link them to e.g. another part of the same or different medium. In the framework of the project we will exploit/extend state-of-the-art technologies in single modality processing and will make a significant step in fusing information from diverse modalities [5], contextual information ([6], [7], [8]) personal context [9], as well as social context. Of great importance will be scalability and granularity issues in the process, focusing on near real-time parallel processing (e.g. using scalable distributed computing software implementations like Hadoop [18]) of extremely large datasets (e.g. the Internet Archive [19] of movies comprising about 45000 MPEG-4 hours and 209287 video items as of October 13th, 2009). Specifically, we foresee contributions in the following fields:

- Effective automatic interactive segmentation (e.g. magic wand, graph or grab cut) of moving objects/regions. It will use current segmentation techniques that require minor user intervention and adapt them to the annotation needs of the ConnectME annotation tool(s).
- Fast real-time tracking of detected objects. Adaptive statistical clustering and feature projection-based classification algorithms will be implemented to identify and track objects that change in appearance through complex and non-stationary background/foreground situations. Objects will be tracked between different parts of the same or different video streams based on visual and semantic similarity.
- Computationally efficient face detection and clustering methods. These technologies will automatically detect faces in large video sequences, cluster them and assign a label to each cluster, which will be extracted from accompanying textual metadata. Similar functionality will be offered for global scenes, which will be identified by text and retrieved based on similarity.
- Large-scale and fast similarity-based retrieval. This will be one of the main contributions of the project, since based on current participants' research novel methods will be proposed and developed that will outperform the state-of-the-art. These methods will be used for all tasks involving similarity-based matching including same and cross media searching (see also the previous contribution).
- Semi-automatic annotation of visual content in broad domains (as opposed to the state-of-the-art which is mostly limited to narrow domains), using knowledge about domain hierarchy, contextual information and any available metadata, for producing rich interpretation of the annotated visual content.
- 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.

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 ConnectME 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, ConnectME 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 nor any kind of social information.

All in all, ConnectME 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 rich media Web and will aid the mainstream trend of producing easily searchable video content.

1.4.4 Web data mining

Web mining has a vital role in meeting ConnectME's objective of utilizing large scale information. Out of the vast amount of content possibly relevant to audiovideo broadcasting, which is available on the Web, only a fraction is available in a semantic format.

Current approaches for using Web content 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). In ConnectME, we aim at bringing the exploitation of both unstructured and structured Web content to a next level by applying Web mining techniques and mapping their results to semantic Web standards. Two types of Web mining will be performed: Web content mining and Web usage mining.

Web content mining identifies relevant Web pages and extracts information from their textual content. 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 that can operate on fully unstructured content[25]. ConnectME will use adaptive information extraction methods (e.g. the ontology-based information extraction tool Ex [56]), which can be trained to extract specific pieces of information from resources of heterogeneous structure. Unstructured (free-text) content will be harnessed with a combination of text mining and natural language processing methods such as THD/SCM [57].

The idea of using results 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 [26] or Boemie [27]. ConnectME 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 ConnectME will utilize **Web usage mining** that analyzes server logs and tracks user behaviour on the Web site. Part of the interaction with the ConnectME platform, which will take place in a Web browser, will be subject to Web usage mining contributing to the personalization experience offered by the platform.

Information mined from the Web can be utilized in ConnectME 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 vbe collected and filtered so that the viewer can be presented information relevant to the current storyline of a shot, scene or the whole broadcast. ConnectME'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). Consistently with the ConnectME's objective to reuse existing technologies where one does not need to reinvent the wheel, we propose building 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[6]. 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.

1.4.5 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 [28] 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) [29]. In the field of segment-based image annotation, several research approaches adopt part-based visual features, either grids ([30]) or

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

segmented regions ([31]), 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 [32], where visual categorization is achieved using a bag-of-keypoints approach, or in [33], 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 [34], 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 [35] 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) [36]. Different approaches for image-level annotation in [37] 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 the Semantic Web language 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 [38]. 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 [39]. 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 [40]. In 2007, Arndt et al. have proposed COMM, the Core Ontology of MultiMedia for annotation [41]. These four ontologies have also been compared [42].

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 [43] 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 [44] 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 [46] 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 "Open Linked Data"[7]. 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). 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 selected concept, reusing the existing and growing Open Linked 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 [47] 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 [48,49,50]. Linking within multimedia presentations, within and among linear and non-linear multimedia presentations is discussed in [50]. [48] discusses links in time-based presentations and proposes a method to specify the

[7] http://linkeddata.org

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 Langauge (SMIL) [51], 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 [52] 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 [53]. 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 ConnectME we will pursue the work of a standardised media fragment addressing scheme which will be supported in the ConnectME platform, as part of the W3C Media Fragments Working Group[8].

1.4.6 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

[8] http://www.w3.org/2008/01/media-fragments-wg.html. The group has published a Working Draft on Fragments at http://www.w3.org/TR/media-frags-reqs/. ConnectME would contribute to a Media Fragments version 2, work on which is expected to be chartered to begin in 2010.

filter a news broadcast down to the stories of interest to the viewer). In ConnectME, 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 [54, 55]. 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 [56].

The ConnectME 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. ConnectME's contribution on user profiling would extend 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.

Behavioural tracking

User behavior and focus of attention tracking will be introduced as additional facet enriching the user profile, beyond explicit preferences introduced by the user, or preferences based on usage put in context. Finer grained user behavior analysis will allow the extension of preference collection and hence user profiles to the level of video frames and objects. On one side, detecting predefined behavioral patterns indicating strong expressed emotions will be proposed. On the other side, computational models of attention will be used at two levels here. First, saliency of particular behavioral pattern will be extracted using models allowing estimation of the rarity of tracked behavioral features. Secondly, the user visual attention profile together with the above behavior interpretation will allow the development of computational models of the particular user attention, in relation with the low-level features of the media file frames and objects, as well the extracted concepts related to these objects.

1.4.7 User interface

ConnectME aims to develop an intuitive user interface, which can be used to access concepts

within audiovisual content and browse information about those concepts in the form of related media. This requires the integration of video content with interactive graphics. This topic has been widely studied in Human-Computer Interaction (HCI) and user interface design. As a result, ConnectME will consider:

- The intuitiveness of the user interface. Video watching is typically a more passive activity carried out at a greater distance from the screen. Hence the interface should be accessible, easy to use and uncomplicated.
- Interaction devices. Video interaction is not necessarily ideal for a point-and-click interface as generally used on a PC in the Web. We will explore more "video-friendly" approaches such as gesture support (for distance viewing) or touchscreen (on a mobile device).

Web video today is shown by the video player application 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. These types of user interfaces are also less suitable when transferred to other non-desktop devices, such as mobile devices or television sets.

In terms of video playout, **ClickVideo**[9] 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. 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 ConnectME, the user interface should be able to have identified hotspots in the display framework itself, rather than to align relevant objects outside a video-player.

[9] http://www.clickvideo.nl/

ConnectME will define an interface layer where interface elements can be defined and overlayed on a video stream using semi-automated media processing and semantic reasoning. Overlays will be W3C standards compliant rather than player proprietary, and will go beyond bounding boxes (current state of the art) to non-trivial object shapes. ConnectME 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.

1.4.8 Multimedia presentation

Generally, the task of assembling different media resources into a meaningful, synchronized presentation today is a manual one, undertaken by 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 semantic-driven interfaces 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 Bordegoni et al. in a Standard Reference Model [62]. Many later systems 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 the 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 that 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 [58]. 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 [59].

More recent work by CWI also relevant to ConnectME, is user exploration of linked data graphs, where different concepts, such as who, what, where and when, can be related to an object in which the user is interested. Exploratory work [60] has shown that while underlying connections among related pieces of information can be used to improve the presentation of potentially related concepts to end-users, that no single interface structure is appropriate for displaying them to end-users. "Who", "what", "where" and "when" are, however, familiar concepts that occur across different domains. Both their familiarity, and the intrinsic spatial and temporal dimensions of "where" and "when", make them amenable to presentation, e.g. of thumbnail images, on maps and timelines [61]. This work will be extended to the background concepts expressed by the active video objects in ConnectME's interface, with particular emphasis on the temporal dimension of video.

This work will be extended, specialized in the topics covered by the ConnectME scenarios and applied to the context of Web-based TV and video in the project.

1.4.9 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 ConnectME. 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

ConnectME has a clear advantage in that many of its consortium members are directly involved in the relevant projects. We provide here comparison to some key projects for ConnectME; a complete list of projects can be found in Appendix I.

Table 3: Projects compared to ConnectME with most relevant projects where results may be adapted and re-used

Project description	Similarities and differences with ConnectME
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/	ConnectME will build on the results from this NoE, since four partners (CERTH, CWI, Eurecom and UEP) already collaborated here. ConnectME has a wider goal than just the focus on media analysis. K-Space did not have specific use cases while ConnectME will further develop analysis and annotation technologies tailored for the needs of Connected Media Experiences and evaluated in two 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. ConnectME will tackle cross-media integration and interaction using near real-time methods that do not rely on large scale offline learning.
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,	ConnectME 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 two specific scenarios. In particular, ConnectME focuses on finer grained annotations of media fragments with individual concepts and on the integration of

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	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 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	MESH is focused on a single application domain (i.e. news) and on the spatial segmentation of news media, while ConnectME develops a platform for generic multimedia services, implements two different scenarios and requires a finer grained spatio-temporal segmentation of different types of media.
Live — Staging of Media Events http://www.ist-live.org	In LIVE (coordinated by Fraunhofer IAIS) 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 ConnectME.
CHORUS — CHORUS is a Coordination Action which aims at creating the conditions of mutual information and cross fertilisation between the projects that will run under Strategic objective 2.6.3 (Advanced search technologies for digital audio-visual content) and beyond the IST initiative. http://www.ist-chorus.org/	Partners of CHORUS (CERTH) participate in ConnectME and therefore the project will benefit from the useful experience and the results of CHORUS. More specifically, ConnectME will reuse the state of the art reports on audio-visual search technologies produced by CHORUS 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 project cluster will be set up when necessary in order to collaborate and re-use existing technology.
INEM4U — Interactive networked experiences in multimedia for you. The aim of iNEM4U is to facilitate enhanced multimedia experiences for individuals and	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

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 is to increase the feeling of "connectedness" with other people or the perception of "involvement" in a particular shared experience. http://www.inem4u.eu/	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 ConnectME, where the intention is to provide multi-device access to content. Within ConnectME more emphasis is given to identifying connection points within media and allowing these to draw in related media assets from multiple networked media sources.
MYMEDIA — Dynamic personalisation of Multimedia	MYMEDIA emphasizes on the personalization of content consumption and recommendation of relevant content to the user. ConnectME 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.
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 ConnectME'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 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.	Many of the results from MultimediaN are relevant to ConnectME. In particular, ConnectME will build directly on results on user 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. ConnectME will go beyond these tasks to investigate how users can navigate through linked video assets. ConnectME 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	Both projects tackle multimedia indexing. ConnectME is more oriented toward the end user (content personalization) with the main goal of developing new experiences of Web-

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	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.
VidiVideo - The Vidi-Video project aims to integrate and develop state of the art components from machine learning, audio event detection, video processing, interaction and visualization into a fully implemented audio-visual search engine. The highlight of the project is the development of a large number of video detectors (in the order of 1000), each corresponding to one concept, thus allowing the association of each video shot with one or more of the approximately 1000 concepts. This will facilitate the concept-based retrieval of video in large collections. http://www.vidi-video.it/	ConnectMe partner CERTH is also a member of the Vidi-Video consortium and actively participates to it, gaining valuable experience in video detector development. As a result of this, ConnectMe will benefit significantly from the developments in Vidi-Video, since the large number of state-of-the-art detectors developed in the latter can serve as an excellent starting point for ConnectMe. However, at the same time there is significant ground to be covered before these detectors can be used towards the goals of ConnectMe. This is mostly due to a) real-time processing being a low priority in Vidi-Video, since the emphasis is on documenting the feasibility and usability of developing a large number of detectors, rather than on solving the problems related to their time-efficiency (i.e. real-time processing requirements) and their applicability to extremely large datasets; b) the detectors of Vidi-Video operating only at the shot level, whereas in ConnectMe the annotation of video not only at shot level but also at the object level is among the major requirements.

We also mention two projects not in the multimedia domain but of relevance to ConnectME. LarKC is an IP working on efficient large scale data processing capabilities on the Web, which we intend to consider as part of the ConnectME platform. OKKAM is a project whose goal is a Linked Entity layer providing unique entity annotation techniques, which can be of value for concept identification in our Linked Media layer.

1.4.10 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 ConnectME would bring to each area of technology.

Table 4: Contribution of ConnectME to relevant technology areas

Area of technology	Current status	Expected status in the next 4 years	ConnectME 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
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 AV material	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)
Personalisation	Possible on a basic level such as video genres, drawing from statically generated user profiling	More refined filtering based on deeper semantic analysis of available data (titles, descriptions...)	Concept-centred filtering based on the object-level annotation of video, incorporation of external behavioural data (e.g. user attention using eye-tracking technology)
User interfaces	Generally “outside” of the video itself, and offering controls	More intuitive interfaces for interactive video,	Interface for the intuitive selection of individual objects in

	related to the video as a whole (e.g. EPG based program descriptions)	enabling more interaction possibilities and eased access to information	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 Connected Media infrastructure in the Web providing global knowledge on concepts and related media, to enable Web-based automatic multimedia presentation generation
Innovative Business Models	Different approaches to monetarizing and little commercial success of current Web TV/Video offers	Commercialization by increased integration of advertising in or beside the video and connection to online shops	Commercialization by way of innovative models that connect Web content to specific objects within the video

1.5 S/T methodology and associated work plan

ConnectME'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 Connected Media parallel to the architectural planning. Both scenarios will also be prepared in terms of storyboards, content acquisition and development, and used subsequently to demonstrate the developed platform.

1.5.1 Architecture

The result of the project will be a platform which comprises all steps for the generation of multimedia presentations composed of meaningfully interconnected Web content. This covers the selection of videos, automatic annotation with meta data based on video speech and graphical objects, linking to external Web sources and other videos and generation of presentations on clients end device.*

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 Web distribution- and service platform for video transfer and user interaction management. The core of the platform will be a novel Connected Media Service Bus, realized on the basis of an APS (Application Server), ESB (Enterprise Service Bus) and WebServices to connect all internal and external components, which allows deploying flexibly other commercial or open APS for the later exploitation. In the project the platform will be implemented and demonstrated using the Thomson APS and ESB.

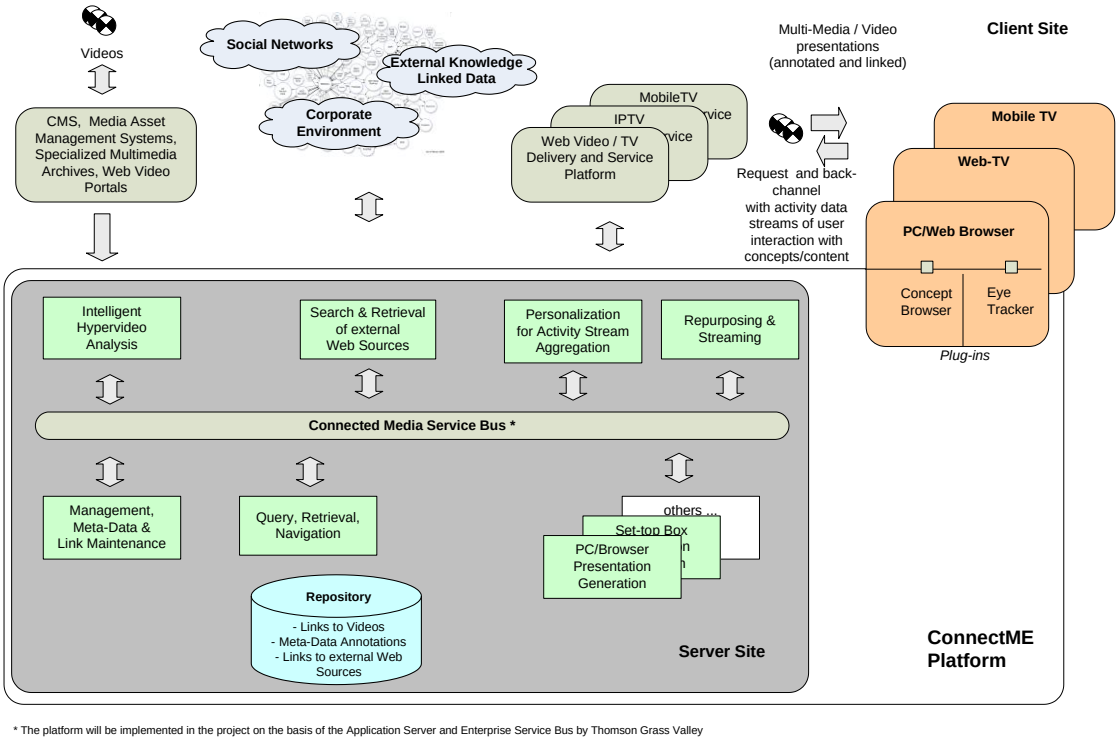
The project intends to deploy the LarKC platform for the reasoning over large amounts of metadata. This platform is currently being developed in the LarKC project which has already provided promising results. But since a final assessment of the benefits from LarKC will not be available during the first 12 months of ConnectME, we will also take the use of scalable RDF triple storages such as the Linda system[10] under consideration.

Specific service components developed within project will be integrated and managed by means of the Connected Media Service Bus. This also holds for the Large Scalable Connected Media Repository which must be able to store, update and retrieve billions of multimedia annotations represented by RDF triples. This will be realized on base of an Open Source high performance RDF Repository and Reasoning Engine; we consider OWLIM as currently first candidate. The main service components comprise Intelligent Hypervideo Analysis, Search & Retrieval of external Web Sources and Knowledge Generation & Linking, a Management Component for managing the CM Service Bus, the Metadata- and Link Maintenance, a Component for delivering the Query, Retrieval, and Navigation for the End User, a Repurposing & Streaming Component, a Personalization component for activity stream aggregation and processing, and a component for the Presentation Generation for different channels, such as normal PC Web Browser, Mobile Devices, and TV Set-Top Boxes.

** we will focus on prototyping innovation user interfaces on PC, using technologies and interaction models which make them transferable outside the project to other devices such as smartphones and TV sets.*

The ConnectME platform is illustrated below:

[10] www.w3.org/2006/02/DERI.pdf



* The platform will be implemented in the project on the basis of the Application Server and Enterprise Service Bus by Thomson Grass Valley

Figure 13. ConnectME platform (high level view)

1.5.2 Content side (broadcasters, archive holders, content owners)

At the broadcaster end, broadcasters can re-use their video production and preparation tools (DW). 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 and provided as a standalone system to the broadcaster use case:

- 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 (Thomson Grass Valley).

1.5.3 Linked media: a supporting Internet-scale information layer

The Connected Media infrastructure is foreseen as including a decentralised, open Web-based information layer formed of 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 Connected Media metadata (media annotation with concepts) will be created and linked in the Web to existing Linked Data metadata (concept metadata and association). The creation of the Linked Media layer will be a project-backed initiative led by WP2 and the partners EURECOM and UEP.

1.5.4 Service side (delivery and network management)

For efficiency and Quality of Service, we will most likely cache Connected Media and Linked Data metadata 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-value data, such as concept viewing and browsing interface, content presentation templates and pre-prepared media presentations as well as aggregrating and processing user activity and behaviour information for service personalisation).

We will develop

- a tool for supporting Connected 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 Connected Media metadata for efficient retrieval (EURECOM/UEP)
- programs for filtering/selecting concepts and related media on the basis of an user profile, generated by their activity (prior interactions with concepts and content) and behaviour (e.g. concepts capturing and holding user attention, measured by eye tracking) (UMONS/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 ConnectME services, browse active concepts, and view sets of media associated with a concept in an appropriate and meaningful way. (CWI)

1.5.5 Web content delivery platform

A Web delivery network for media content and evaluation infrastructure (testbeds) will be provided by CONDAT. It will:

- Organize parallel media streams,
- Manage the added-value data services,
- Configure networks,
- Configure subscribers (incl. user profiles),
- Monitor services.

Thomson Grass Valley will provide an application server with an ESB. This platform will host modules developed either by Thomson or by other 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 networks 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.

1.5.6 End user device

The core purpose of end device development will be to provide an open, multi-layered, client and server-customisable device front-end on which all ConnectME media and content can be projected onto, and which will be called upon over open Web standards, e.g. using a REST API. In order to project available media, it needs to interlink to a variety of metadata descriptors. It will be able to interface with the services in ConnectME and will need to be able to interface to third party non-ConnectME 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 interface description, allowing smooth integration with third party XML-driven services.

We will consider the use of HTML, SMIL, Ajax "widgets" and Flash in a ConnectME player which will connect to the variety of services developed within the ConnectME project and will provide innovative in-video interaction with on-screen objects. For implementation of the ConnectME front end on an interactive video player CWI can take the AMBULANT video player[11] as a basis, which is compliant with open standards such as W3C SMIL 3.0.

ConnectME 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. 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 selection mechanism will enable end-users to navigate to other related content in the video itself or other available media. An API will be made available which will interact with the ConnectME platform.

Work Packages

We organize the work in workpackages so that the research and development of the ConnectME 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 making Web-based connected media possible**: WP1 will focus on challenges of scalable concept-level processing 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 **Connected Media Experience**: the interactive access to concepts and content through an intuitive user interface and the dynamic generation of multimedia presentations to present information packages related to the selected concept will be tackled in WP3; while the filtering and personalisation of the concept browsing and content presentation to the user according to their preferences and attention which will be captured by innovative activity and behaviour tracking approaches and formalised 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 **ConnectME 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 Web media delivery capabilities and developing the front end application with interactive video and concept/content browsing capabilities. Applying this in use case **scenarios** (WP6) will be the means to validate the ConnectME platform functionality and achieve its transfer – prototypically – into actual Web TV and video services. The scenarios will involve a test environment to deliver ConnectME 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.

[11] http://www.ambulantplayer.org/

Finally, the prototypical use of the ConnectME 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 ConnectME workpackages:

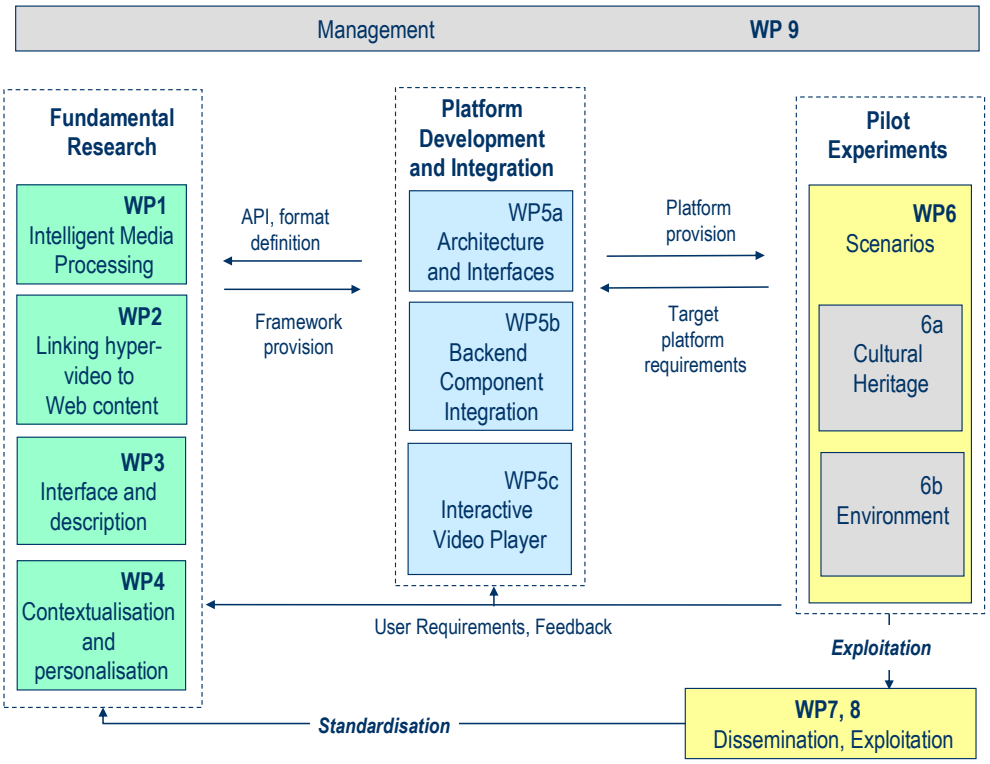


Figure 14: WP structure

Detailed workplan description

Work package 1: Intelligent media processing

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 of video sequences. It will offer the technologies to detect objects/regions of interest in a computationally efficient way with respect to both large scale datasets and timing, track them over time, identify them using textual metadata and finally link them to e.g. another part of the same or other video or to external sources. Overall, it addresses visual regions detection, objects and scene labeling, as well as large-scale matching and retrieval of visual information as a prerequisite for enabling association between regions detected in the stream or/and other form of information (e.g. text, web links etc.).
The work herein addresses also the computational complexity and scalability issues that arise from the desired application of advanced video analysis techniques to large video datasets. More specifically, by focusing the development work on algorithms and techniques ran on

cloud computing infrastructure that can support near real-time processing, it becomes highly scalable and suitable for use in large-scale applications.
Ultimately, this WP will provide functionalities both to media annotators and to end-users; human annotators will be provided with selected visual regions of interest (either spatial or spatiotemporal) so as to minimize the time for manual refinement and ease the content labeling process. Based on similarity, instances of the annotated regions found in the same stream will be also labeled accordingly. The user will then be able to select the previously annotated regions, search for similar ones in the collection and receive higher level information about them (e.g. web abstracts).

T1.1 Visual segment localization and tracking

This task deals with the processing and extraction methods involved in the first-stage analysis of the visual data that will aid the semi-automatic annotation of the content. The outcome will mostly apply to the authoring part of the process. Specifically, this task will start with implementing a real-time technique for video segmentation to shots, adopting the most appropriate technique of the relevant literature in terms of accuracy and time-efficiency for use in the ConnectME system. The task will subsequently develop and implement real-time or near-real-time techniques for region of interest detection and for video segmentation to spatiotemporal objects (i.e., region of interest detection and tracking). Real-time compressed domain techniques will be investigated to this end, including but not limited to techniques for generating fine-grained (i.e. at near-pixel level, rather than at block/macroblock level) approximate color and texture information without the need for video decompression. Regions of interest will be detected either interactively (i.e. combining the video processing techniques with human input, e.g. clicking or drawing a line over an object of interest) or fully automatically by detecting regions/features that exhibit salient behavior, e.g. consistent motion. In the development of the techniques of this task, equal emphasis will be put on optimizing the accuracy and the scalability and computational complexity of the developed techniques, with real-time processing (or near-real-time, when it comes to semi-automatic techniques, since human interaction is inevitably non-real-time) being the objective.

T1.2 Object & scene labeling

Based on the outcome of T1.1 the stream will be decomposed into regions enclosing objects or generic regions of interest. The latter will be detected using techniques such as time- and memory-efficient clustering based on descriptors' similarity, or trained concept detectors based on state-of-the-art machine learning techniques. The developed clustering and concept detection techniques will be based to the greatest extent possible on compressed video information (e.g. color, motion, texture) also exploited in T1.1, to minimize computational complexity. Detectors will be developed not only for shots but also for content items at the finer levels of granularity supported by the segment localization and tracking techniques of T1.1, e.g. for spatiotemporal objects. With clustering and learning-based techniques typically being particularly computationally intensive, work will also be carried out in this task on exploiting currently unexploited hardware resources (e.g. Graphics Processing Units) and cloud computing (e.g. Amazon EC2) to the maximum extent possible, to achieve real-time or near-real-time processing on non-specialized hardware, consequently ensuring the high scalability of the developed techniques.

This task will result in the acquisition of semantic descriptions of objects and/or scenes. The same functionality will be offered for global scenes, whereas similar instances will be detected by clustering based on descriptors' similarity. Focus will be given on objects accompanied by rich metadata information, such as external web sources of specific scope like IMDB or Wikipedia. Of special interest are efficient face detection methods (real-time identification of regions depicting faces) and the exploitation of accompanying textual metadata, such as available scripts or subtitles to label them by e.g. actor names publicly available e.g. in IMDB. For example, the user will be able to click on an unknown face and

let the method identify it through the script, detect similar instances in the same or different video sequence and finally link it to available external information sources (e.g. web or blog info). Similarly, the user will be able to select a global scene (e.g. Eiffel tower by night) and be provided with related web info or similar scenes in the same or other media.

T1.3 Similarity-based retrieval

This task will mainly propose and implement methods for matching and retrieval based on visual similarity. Appropriate visual features and descriptors will be detected and extracted respectively in order to efficiently represent the input. Large scale matching techniques will then be applied, mostly aimed to provide fast and accurate retrieval of similar scenes/objects. Emphasis will be put on computational efficiency so as to confront the huge amount of data to become part of the project's collection.

T1.4 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).

T1.5 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 ConnectME service providers. The tool will use the combination of analysis approaches developed in earlier Tasks of the Work Package to automatically perform a number of analysis tasks and at the same time allow for user interaction, by e.g. making informed suggestions to users of objects of interest and their label. For example, during the analysis the user will be able to click on an unknown face and let the automatic analysis methods to identify it by analyzing the associated script; detect similar instances in the same or different video sequence using different features; and finally link it to available external information sources (e.g. web or blog info). Using this tool the users should be able to complete the annotation process with a minimal effort and in minimal time. As a basis, partners can provide existing tools from the FP6 projects aceMedia and LIVE for extension and improvement.

T1.6 Intelligent hypervideo analysis requirements and evaluation

This task refers to all activities relating to the definition of requirements for content analysis, to be performed during the first six months of the project, and the subsequent continuous technical evaluation of the content analysis techniques being developed in tasks T1.1-T1.5 against these requirements. The latter evaluation activities include setup of ground truth data, manual or semi-automatic annotation of available content according to the requirements of each task, design of evaluation metrics and methodologies, and performance of experiments to measure the effectiveness of the developed technologies. It does not refer to collection of content itself, nor to the development of content annotator tools. Particular focus will be given in validating that the WP1 outcome outperforms conventional approaches in the field. In line with the objectives set for the development of the analysis methods in the Work Package, the effectiveness of the developed technologies will be assessed with respect to both accuracy of

results and computational complexity / scalability, to ensure that they are both highly accurate and highly efficient when applied to extremely large datasets.

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 Open 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 T2.1 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 T2.2 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 ConnectME use cases. We will implement an API and the necessary interfaces to be integrated in the ConnectME 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 video 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 video material, 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: ConnectME 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 ConnectME 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 ConnectME 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. making cultural heritage accessible), or information gathering for professional information users (e.g. environmental facts and statistics). 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 both 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 ConnectME 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 ConnectME platform will be carried out. This will be done with potential users of the ConnectME 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

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 as webcams or microphones on user behavior including the focus of attention (gaze detection, quantity of motion, contraction index, symmetry index, voice signals, etc…). 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 the rarity index variation of the current user with 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, the mean of their behavior rarity index can provide a mean emotional evolution profile. In this 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.

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: ConnectME platform

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

The work package has the following objectives:

- To define an architecture that integrates the components for video/audio and Web content analysis, metadata generation and storage, creation of interactive video and user interfaces, and their delivery and presentation to the user into an end-to-end platform
- To define APIs and protocols to connect all components and data flows.
- To implement interfaces, data transfer and input/output functions not provided by WPs 1-4
- To integrate and test of all components designed and implemented in WPs 1-4
- To coordinate 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.
- To develop the front-end system which provides the interactive video player and ConnectME user interface, ensuring interoperability with the back-end data and processes
- To coordinate and self-assess 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. The core result will be the *Connected Media Service Bus*.

This core component will be built by using the Thomson Application Server and Thomson Enterprise Service Bus as integration platform, and delivery modules will be provided for video and data delivery over IP networks.

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 3rd 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 flexibly 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 co-operation 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 interactive video player

The front-end playout will be enabled by the development of a interactive video player which enables the layering over video of further interactive regions allowing for the presentation of the ConnectME 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 presentation layer.

The front-end needs to be interactive and fast and therefore it will be providing ConnectME 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 ConnectME platform provided in WP5. Finally, evaluations of the scenarios will be organized through internal testing, user trials and public feedback gathering.

6a: 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.

6b: Environment

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 ConnectME consortium selected a scenario dealing with environmental and economic issues as it can be certain that these topics will remain on the news (and world) 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 / economic field, the system and underlying technology can easily be ported to other domains.

An important aspect of the WP6 activity will be the user trials and evaluation (T6.3)

The first user trials will be conducted by involved employees of the partners. Selected employees will get access to the ConnectME training material and environment in a controlled experimental environment. The evaluation will be performed based on observation how users use the system, questionnaires for the users and focused groups.

The second trial will go beyond an experimental environment and will provide selected employees with access to ConnectME within their everyday routines. Selected employees, mainly such that have tasks that might take advantage of the ConnectME environment, will get access to ConnectME and can use it within their daily routines. The evaluation will be performed based on observation, questionnaires for the users and focused groups.

The final trail will involve potential independent users and crowd sourcing. To achieve that, a special test site will be created, which will provide training material for interested users as well as access to ConnectME functionalities. Within a ConnectME competition independent users will be motivated to use ConnectME and to solve specific tasks related to the usage scenarios by using the ConnectME technology. The evaluation will involve crowd sourcing, analysis of the results achieved by the users, and online surveys.

Work package 7: Dissemination

One of the aims of the ConnectME 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 ConnectME activities are disseminated to various communities and interest groups. These include potential end customers of future ConnectME 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 ConnectME 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 ConnectME 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 ConnectME, 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 ConnectME will have the best possible impact both in the scientific and commercial communities. In terms of current standardization efforts, ConnectME 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 ConnectME can have an impact, by providing a market analysis, developing effective business models for ConnectME services and exploitation strategies for ConnectME 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 Web TV and video market based on a broad market analysis. A second goal of this WP is assessment of user requirements based on qualitative and quantitive 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 providers, Web TV and video portals 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

ConnectME will develop an innovative platform that has the potential to change considerably how video and TV on the Web 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 interlinking of content is 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 regarding which part of the content to connect? How does ConnectME 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 ConnectME? Finally, an open question is if and in which form advertising might get included in ConnectMe 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 providers and Web TV/video portals as well as advertisers.

Activities

- **Overall market analysis** about market potential and trends based on 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 ConnectME.
- **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 providers and Web video/TV portals. 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 of ConnectME has the potential to change existing content and advertising value chains for TV and video on the Web. 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 ConnectME 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 ConnectME 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. ConnectME offers business opportunities for all involved partners in the project and all involved players in the ConnectME value chain. A particular focus will be given to open source business models: potential open source licensing models will be explored together with related business strategies of involved companies. Besides that, potential to create an open source community around the ConnectMe platform will be investigated and incentive structures will already be incorporated in the broad user trial. 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. 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 ConnectME 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

WP No[12]	Workpackage title	Type of activity [13]	Lead partici-pant no[14]	Lead participant short name	Person month-s[15]	Start[16]	End [17]
1	Intelligent media processing	RTD	2	CERTH	123	1	36
2	Linking hypervideo to web content	RTD	6	EURECOM	125	1	36
3	ConnectME interface and presentation	RTD	5	CWI	76	1	36
4	Personalisation and contextualisation	RTD	1	IAIS	124	1	36
5	ConnectME platform	RTD	9	CONDAT	115	1	36
6	Scenarios	RTD/ DEM	8	STI2	113	1	36
7	Dissemination	RTD/ OTHER	8	STI2	61	1	36

[12] Work package number: WP 1 – WP n.

[13] 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.

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

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

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

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

8	Exploitation	RTD	7	USG	65	1	36
9	Management	MGT	1	IAIS	29	1	36
	TOTAL				831	1	36

Table 5: List of Deliverables[18]

Del. No[19]	Deliverable name	WP no.	Nature [20]	Dissemination level[21]	Delivery date[22] (proj. month)
D7.1	Project website	7	R	PU	2
D1.1	State of the art and requirements analysis for hypervideo	1	R	PU	6
D2.1	Specification of the Linked 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	ConnectME 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
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	ConnectME 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	ConnectME front-end: interactive video playback	5	P	PU	12
D6.1	Scenario descriptions	6	R	PU	12
D8.2	First market analysis	8	R	PU	12
D1.2	Hypervideo analysis tools and evaluation v1	1	R	PU	18

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

[19] 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.

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

[21] 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).

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

D5.3	First ConnectME end-to-end platform	5	P	PU	18
D7.3	Dissemination and standardisation report v1	7	R	PU	18
D8.3	ConnectME business models, v1	8	R	PP	18
D2.4	Annotation and retrieval of media by concept in the Web	2	R	PU	24
D2.5	Final ConnectME 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	ConnectME Service Bus	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	ConnectME end-to-end platform value chain	8	R	PU	24
D1.3	Hypervideo analysis tools and evaluation v2	1	R	PU	27
D5.5	ConnectME front-end: interactive video playback v2	5	P	PU	27
D6.3	User trial results	6	R	PU	27
D7.5	Dissemination and standardisation report v2	7	R	PU	27
D2.6	Advanced concept labelling by complementary Web mining	2	R	PU	30
D3.7	ConnectME user interfaces v2	3	R	PU	30
D4.6	Contextualisation solution and implementation	4	R	PU	30
D5.6	Final ConnectME end-to-end platform	5	P	PU	30
D6.4	Scenario demonstrators	6	D	PU	30
D7.6	Project demonstrator v2	7	D	PU	30
D8.6	Market and product survey for ConnectME services and technology	8	R	PU	30
D8.7	ConnectME business models, v2	8	R	PP	30
D1.3	Hypervideo analysis tools and evaluation final version	1	R	PU	36
D2.7	Final Linked Media layer and evaluation	2	R	PU	36
D3.8	Design guideline document for concept-based presentations	3	R	PU	36
D4.7	Evaluation and final results	4	R	PU	36
D5.7	Validation of the ConnectME architecture	5	R	PU	36

D6.5	Final evaluation	6	R	PU	36
D7.7	Dissemination and standardisation report v3	7	R	PU	36
D8.8	Common and individual exploitation plans for after the project	8	R	PP	36

*The ConnectME Service Bus may be subject to IPRs from Thomson Grass Valley. The ConnectME end-to-end platform, made up of the back end and front end components for ConnectME services, will be made available as open source.

WP	WP and Task Descriptions	Start	End	Duration	Year 1												Year 2												Year 3											
		Month	Month	Month	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
					1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
1	Intelligent media processing	1	36	36																																				
T1.1	Visual segment localization and tracking	1	24	24																																				
T1.2	Object and scene labelling	7	30	24																																				
T1.3	Similarity-based retrieval	7	30	24																																				
T1.4	Complementary audio and text analysis	7	24	18																																				
T1.5	User assisted annotation tool	7	30	24																																				
T1.6	Intelligent hypervideo analysis evaluation	31	36	6																																				

WP	WP and Task Descriptions	Start	End	Duration	Year 1														Year 2									Year 3											
		Month	Month	Month	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
					1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
2	Linking hypervideo to web content	1	36	36																																			
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	36	36																																			
T2.4	Retrieving additional content from the web	13	36	24																																			

WP	WP and Task Descriptions	Start	End	Duration	Year 1																			Year 2															Year 3														
		Month	Month	Month	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
					1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
3	ConnectME interface and presentation engine	1	36	36																																															
T3.1	User requirement analysis	1	6	6																																														
T3.2	User-centered design and specification of hypervideo information interfaces	1	30	18																																														
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	36	12																																														

WP	WP and Task Descrip-tions	Start Month	End Month	Duration Month	Year 1												Year 2												Year 3											
					M 1	M 2	M 3	M 4	M 5	M 6	M 7	M 8	M 9	M 10	M 11	M 12	M 13	M 14	M 15	M 16	M 17	M 18	M 19	M 20	M 21	M 22	M 23	M 24	M 25	M 26	M 27	M 28	M 29	M 30	M 31	M 32	M 33	M 34	M 35	M 36
4	Contextualisation and Personalisation	1	36	36																																			
T4.1	User profiling	1	24	24																																			
T4.2	Contextualisation	1	30	30																																			
T4.3	Personalized content filter	7	30	24																																			
T4.4	Evaluation of personalisation technologies	31	36	6																																			

WP	WP and Task Descrip-tions	Start	End	Duration	Year 1																																	
		(Month)	(Month)	(Month)	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
					1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
5	ConnectME platform	1	36	36																																			
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	30	24																																			
T5.4	Player API	7	27	9																																			
T5.5	Interactive Video Playback	7	27	9																																			
T5.6	Front-end development and component integration	7	30	24																																			
T5.7	Validation	31	36	6																																			

WP	WP and Task Descrip-tions	Start	End	Duration	Year 1												Year 2																							
		(Month)	(Month)	(Month)	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
					1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
6	Scenarios	1	36	36																																				
T6.1	Media and content preparation	1	27	9																																				
T6.2	Scenario realization on ConnectME platform	7	30	15																																				
T6.3	User trials and evaluation	13	36	15																																				

WP	WP and Task Descrip-tions	Start Month	End Month	Duration Month	Year 1												Year 2												Year 3											
					M 1	M 2	M 3	M 4	M 5	M 6	M 7	M 8	M 9	M 10	M 11	M 12	M 13	M 14	M 15	M 16	M 17	M 18	M 19	M 20	M 21	M 22	M 23	M 24	M 25	M 26	M 27	M 28	M 29	M 30	M 31	M 32	M 33	M 34	M 35	M 36
7	Dissemination	1	36	36																																			
T7.1	Dissemination and standardisation planning	1	36	12																																			
T7.2	Project website	1	36	36																																			
T7.3	Production of dissemination materials and publications	7	36	30																																			
T7.4	Project presentations at events	7	36	30																																			
T7.5	Project demonstrator	19	36	18																																			
T7.6	Standardisation activities	7	36	30																																			

WP	WP and Task Descrip-tions	Start	End	Duration	Year 1										Year 2										Year 3													
		(Month)	(Month)	(Month)	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
					1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
8	Markets, Businnes Models and Exploitation Strategies	1	36	36																							h											
T8.1	Product & market surveys	1	30	24																																		
T8.2	Business models	7	30	24																																		
T8.3	Exploitation planning & Exploitation activities	1	36	36																																		

WP	WP and Task Descrip-tions	Start	End	Duration	Year 1										Year 2										Year 3													
		(Month)	(Month)	(Month)	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
9					**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**
T9.1	Project Management	1	36	36																																		
T9.2	Periodic Scientific Reports	1	36	36																																		
T9.3	Periodic Financial Reports	1	36	36																																		
T9.4	Work self-assessment and workplan revision	7	36	18																																		
T9.5	Internal communication structures	1	36	36																																		

Work package descriptions

Workpackage number	1	**Start date or starting event:**			M1		
Workpackage title: Intelligent media processing							
Activity type[23]: RTD							
Participant id	1	**2**	3	4	10		
Person-months per participant:	36	38	18	15	16		

Objectives

- Automatically detect and track objects or regions of interest in audiovisual material in order to support the annotation process.
- Automatically perform object and scene labelling, resulting in fine-grained semantic descriptions of scenes
- Exploit compressed video information and currently unexploited hardware resources, to enable analysis in real-time, thus satisfying very high scalability requirements
- Provide an intelligent annotation tool which reduces the manual effort required to annotate material with concepts

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

T1.1 Visual segment localization and tracking (CERTH-ITI, IAIS, UMONS) (M1-24)

T1.1.1: Automatic visual analysis techniques (M1-24)

This task will provide methods for (semi-)automatic annotation of audiovisual content, such as video segmentation to shots or spatiotemporal objects.

T1.1.2: Interactive visual analysis techniques (M7-24)

This sub-task will focus on detecting regions of interest, either fully interactively (i.e. by combining video processing techniques with human input) or fully automatically (i.e. by detecting regions/features that exhibit salient behaviour).

T1.2 Object and scene labelling (CERTH-ITI, UEP) (M7-30)

T1.2.1: Clustering and learning-based techniques for labelling (M7-27)

This sub-task will use techniques such as time- and memory-efficient clustering based on descriptors' similarity or state-of-the-art machine learning techniques to detect regions of interest and will result in the acquisition of semantic descriptions for these regions.

T1.2.2: Software and hardware optimization of labelling (M18-30)

This sub-task will exploit currently unexploited software and hardware resources to achieve real-time or near-real-time content processing and ensure high scalability of developed techniques.

T1.3 Similarity-based retrieval (CERTH-ITI, IAIS) (M7-30)

T1.3.1: Retrieval of common visual features and descriptors (M7-12)

This task will mainly propose and implement methods for matching and retrieval based on visual similarity by detecting and extracting appropriate visual features and descriptors.

[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.

T1.3.2: Extensions for more complex visual material (M13-24)

This sub-task will utilize the results of previous sub-task to extend and apply them to more complex audiovisual content.

T1.3.3: Scalable algorithms for retrieval across large collections (M25-30)

This sub-task will apply large scale content matching techniques aimed to provide fast and accurate retrieval of similar scenes/objects.

T1.4 Complementary audio and text analysis (S+V, CERTH-ITI, UEP) (M1-24)

T1.4.1: Audio analysis techniques (M1-24)

This sub-task will deploy audio analysis technologies to generate multimedia object segmentations based on the audio stream.

T1.4.2: Text analysis techniques (M1-24)

This sub-task will analyse and associate available collateral textual information sources directly to the audiovisual content.

T1.5 User assisted annotation tool (IAIS, CERTH-ITI) (M7-30)

T1.5.1: Review of existing annotation tools (M7-12)

This sub-task will review, extend and improve existing tools from related FP6 projects, such as aceMedia or LIVE.

T1.5.2: ConnectMe annotation tool (M12-30)

This sub-task will provide a prototype of the ConnectMe annotation tool that will be implemented to demonstrate the annotation process.

T1.6 Intelligent hypervideo analysis requirements and evaluation (CERTH-ITI, S+V, IAIS) (M31-36)

T1.6.1: Requirements analysis (M31-32)

This sub-task will deal with issues handling setup of ground truth data and manual or semi-automatic annotation of available content according to the project's requirements.

T1.6.2: Evaluation of techniques and tools (M33-36)

This sub-task will evaluate techniques and tools of the intelligent hypervideo processing of Tasks T1.1-1.5.

Deliverables (brief description) and month of delivery

D1.1 State of the art and requirements analysis for hypervideo (M6)

D1.2 Hypervideo analysis tools and evaluation, first version (M18)

D1.3 Hypervideo analysis tools and evaluation, second version (M27)

D1.4 Hypervideo analysis tools and evaluation, final version (M36)

Milestones and expected result

(M12) Implementation of hypervideo annotation tool v1 and core hypervideo analysis functionality
(M24) Implementation of hypervideo annotation tool v2 and refined hypervideo analysis functionality

(M30) Implementation of hypervideo annotation tool v3 and final hypervideo analysis functionality

(M36) Final results and evaluation

Workpackage number	2	**Start date or starting event:**		M1			
Workpackage title: Linking hypervideo to web content							
Activity type: RTD							
Participant id	2	3	**6**	9			
Person-months per participant:	12	48	51	14			

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 Linked Media layer.
- Develop the infrastructure to crawl relevant parts of the Web and mine concepts, creating and searching the Linked 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)(M1-36)

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 (M7-30)

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 (M13-36)

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)(M13-36)

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

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 (M25-36)

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.

T2.4.3 Web mining approaches to annotating Web-based content and gathering information around concepts (M13-36)

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 (M30)
D2.7 Final Linked Media layer and evaluation (M36)

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 Connected Media Layer
(M36) Multimedia metadata API (final specification), Web mining for concept identification and disambiguation (final specification), Evaluation and final results

Workpackage number	3	**Start date or starting event:**		M1		
Workpackage title: ConnectME interface and presentation engine						
Activity type: RTD						
Participant id	**5**	8	10	11		
Person-months per participant:	52	6	9	9		

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 (CWI, S+V, DW)**(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 two 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)**(M1-12, M25-30)
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 (M25-30)
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 (CWI)** (M7-24)
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 ConnectME platform.
T3.3.2 Refinement of the interface and presentation engine (M13-24)
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, DW)**(M19-24, M31-36)
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 (M31-36)
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)
D3.3 ConnectME user interfaces sketch (M12)
D3.4 ConnectME interface and presentation engine version 1 (M12)
D3.5 Requirements document for ConnectME user interfaces (version 2) (M24)
D3.6 ConnectME interface and presentation engine version 2 (M24)
D3.7 ConnectME user interfaces selected and refined (version 2) (M30)
D3.8 Design guideline document for concept-based presentations (M36)

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 ConnectME platform.
M24 Refined presentation engine.
M30 Final interfaces decided and implemented.
M36 Final evaluation of interfaces in the user trials.

Workpackage number	4	**Start date or starting event:**			M1		
Workpackage title: Contextualisation and personalisation							
Activity type: RTD							
Participant id	**1**	2	3	4	5	8	
Person-months per participant:	34	28	20	27	8	4	

Objectives

- Specify a schema for user profiles
- Specify how to capture user profiles from activity graphs – how users interact with content
- Specify how to capture user profiles from behavioural tracking – eye and body movement
- 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 tracked behaviour

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

T4.1 User profiling (IAIS, CERTH-ITI, 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 Contextualisation (UMONS, CERTH-ITI) (M1-M30)

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 (M25-30)
This subtask will provide the mechanism to expand the personalisation of the ConnectME 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 (IAIS, CERTH-ITI, UEP, UMON) (M7-M30)
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 (M25-30)
Final implementation of the profile-content semantic matcher, incorporating advanced concept and content filtering.

T4.4 Evaluation of personalization technologies (IAIS, CERTH-ITI, UMON) (M31-M36)
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 (M30)
D4.6 Evaluation and final results (M36)

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
(M30) Final specifications for content filtering/adaptation, final personalization filter component
(M36) Final results and evaluation

Workpackage number	5		**Start date or starting event:**		M1	
Workpackage title: ConnectME platform						
Activity type: RTD						
Participant id	2	5	6	**9**	12	
Person-months per participant:	4	21	4	44	42	

Objectives

This WP realises the integration of components of ConnectME 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 (Linked Media layer)
- interlinking of 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

It will integrate or provide the following components on the front-end:

- interactive video playback
- user interface layer for the ConnectME services (the Player API)
- user activity/behaviour 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 (TGV)(M1-24)
T5.1.1 Specification and set-up of the integrating platform (ConnectME Service Bus) (M1-6)
T5.1.2 Development of delivery modules for Web TV and video (M7-12)
T5.1.3 Testing of delivery of Web TV and video (M13-18)
T5.1.4 Integration platform testing and reconfiguration (M19-24)

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

T5.3 Back-end development and component integration (CONDAT, TGV)(M7-30)
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 (M25-27)
T5.3.5 Final refinement of back-end system (M28-30)

T5.4 Player API (CWI)(M7-12, 25-27)
T5.4.1 First API and display tool (M7-12)
T5.4.2 Second API and display tool (M25-27)

T5.5 Interactive Video Playback (CWI, EURECOM)(M7-12, 25-27)
T5.5.1 First Interactive Video player (M7-12)
T5.5.2 Second Interactive Video player (M25-27)

T5.6 Front-end development and component integration (CWI, EURECOM, CERTH-ITI)(M7-30)
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 (M19-27)
T5.6.4 Final refinement of back-end system (M28-30)

T5.7 Validation (CONDAT, TGV, CWI)(M31-36)
T5.7.1 Validation of the front-end (M31-36)
T5.7.2 Validation of the back-end (M31-36)
T5.7.3 Validation of the end-to-end platform (M31-36)

Deliverables (brief description) and month of delivery

D5.1 ConnectME platform and architecture (M6)
D5.2 ConnectME front end: interactive video player and API (M12)
D5.3 First ConnectME end-to-end platform (M18)
D5.4 ConnectME Service Bus (M24)
D5.5 ConnectME front end: interactive video player and API V2 (M27)
D5.6 Final ConnectME end-to-end platform (M30)
D5.7 Validation of the ConnectME architecture (M36)

Milestones and expected result

(M6) First specifications of the ConnectME 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 video
(M24) Final specification of the integrating platform (the ConnectME Service Bus)
(M27) Final development/deployment of components in the back-end and front-end
(M30) Final implementation of the end-to-end platform
(M36) Validation of the end-to-end platform

Workpackage number	6		**Start date or starting event:**			M1			
Workpackage title: Scenarios									
Activity type: RTD									
Participant id	1	2	3	6	7	**8**	10	11	12	
Person-months per participant:	2	2	4	6	7	14	36	36	6	

Objectives

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

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

T6.1 Media and content preparation (STI, DW, S+V, CERTH-ITI, UEP, EURECOM)(M1-M27)
T6.1.1 Initial media and content preparation (M1-6)
T6.1.2 Revised media and content preparation (M25-27)

T6.2 Scenario realization on ConnectME platform (STI, DW, S+V, IAIS)(M7-M30)
T6.2.1 Storyboarding of the scenarios (M7-12)
T6.2.2 Scenario realization on the ConnectME platform (M19-24)
T6.2.3 Revised scenario realization on the ConnectME platform (M28-30)

T6.3 User trials and evaluation (STI, DW, S+V, UNISG)(M13-M36)
T6.3.1 Internal trials with the partners (M13-18)
T6.3.2 User trials held by the partners (M25-27)
T6.3.3 Final user trials held by the partners (M31-36)

Deliverables (brief description) and month of delivery
D6.1 Scenario descriptions (M12)
D6.2 Scenario demonstrators (M24)
D6.3 User trial results (M27)
D6.4 Scenario demonstrators (M30)
D6.5 Final evaluation (M36)

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 ConnectME platform and demonstrators are available
(M27) User trials have taken place arranged by the partners
(M30) Scenarios have been refined and realized anew on the ConnectME platform
(M36) Final user trials have taken place and the scenarios have been evaluated

Workpackage number	7			**Start date or starting event:**					M1		
Workpackage title: Dissemination											
Activity type: RTD/OTH											
Participant id	1	2	3	4	5	**6**	7	8	9	10	11	12
Person-months per participant:	6	3	3	3	3	13	3	18	2	2	6	2

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 ConnectME 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, all)(M1-6, 17-18, 29-30, 35-36)

T7.2 Project website (STI)(M1-36)

T7.3 Production of dissemination materials and publications (STI, all)(M7-36)

T7.4 Project presentations at events (STI, all)(M7-36)

T7.5 Project demonstrator (STI, all)(M19-36)
T7.5.1 First project demonstrator (M19-24)
T7.5.2 Second project demonstrator (M25-30)
T7.5.3 Dissemination of the demonstrator online (M31-36)

T7.6 Standardisation activities (EURECOM, all)(M7-36)
T7.6.1 First standardisation push (M7-18)
T7.6.2 Evaluation of standardisation efforts (M19-24)
T7.6.3 Second standardisation push (M25-30)
T7.6.4 Sustainability actions for standardisation activities after the project (M31-36)

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 (M30)
D7.7 Dissemination and standardisation report v3 (M36)

Milestones and expected result
(M6) Plan for dissemination and standardisation during the project is fixed, project website is online and 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) A final demonstrator is ready for presenting the results of the project , another push for standardisation of project results
(M36) 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: RTD										
Participant id	1	**7**	8	9	10	11	12			
Person-months per participant:	9	28	10	6	6	4	2			

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 ConnectME 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 ConnectME results can be turned in commercial success

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

T8.1 Product & Market Surveys (UNISG, CONDAT, S+V, DW, TGV)(M1-M30)

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

The goal of this first market analysis is to explore how Web video and TV is used by users today and what their preferences for interactive Web video and TV 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 ConnectME technology is designed. The concrete concepts will be evaluated by the user. The goaö will be to evaluate the potential adoption of ConnectMe 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 (M25-27)

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

T8.1.5 Overall market analysis for exploitation of ConnectME results (M28-36)

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, IAIS)(M7-M30)

T8.2.1 Business model development (M7-18, 25-30)

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 ConnectME 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, STI, CONDAT, S+V, DW, TGV)(M1-M36)

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 (M31-36)

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 ConnectME business models, v1 (M18)

D8.4 Evaluation of exploitation impact and recommendations (M24)

D8.5 ConnectME end-to-end platform value chain (M24)

D8.6 Market and product survey for ConnectME services and technology (M30)

D8.7 ConnectME business models, v2 (M30)

D8.8 Common and individual exploitation plans for after the project (M36)

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 ConnectME value chain has been completed

(M30) final market analysis, final business plans and final exploitation activities are complete

(M36) common and individual exploitation plans for ConnectME results have been produced

Workpackage number	9		**Start date or starting event:**			M1					
Workpackage title: Management											
Activity type: MGMT											
Participant id	**1**	2	3	4	5	6	7	8	9	10	11	12
Person-months per participant:	18	1	1	1	1	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 (M1-36) (**IAIS**)*
*T9.2 Periodic Scientific Reports (M1-36) **(IAIS, all)***
*T9.3 Periodic Financial Reports (M1-36) (**IAIS**, all)*
*T9.4 Work Self-assessment and Workplan Revision (M7-12, 19-24, 31-36) (**IAIS**, all)*
*T9.5 Internal Communication Structures: mailing lists, wiki etc. (M1-36) (**IAIS**)*

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 (M36)

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
(M36) Final report with summary of project results and sustainability actions

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	*IAIS*	*36*			*34*		*2*	*6*	*9*	*18*	*105*
2	*CERTH*	*38*	*12*		*28*	*4*	*2*	*3*		*1*	*88*
3	*UEP*	*18*	*48*		*20*		*4*	*3*		*1*	*94*
4	*UMONS*	*15*			*27*			*3*		*1*	*46*
5	*CWI*			*52*	*8*	*21*		*3*		*1*	*85*
6	*EURECOM*		*51*			*4*	*6*	*13*		*1*	*75*
7	*USG*						*7*	*3*	*28*	*1*	*39*
8	*STI*			*6*	*4*		*14*	*18*	*10*	*1*	*53*
9	*CONDAT*		*14*			*44*		*2*	*6*	*1*	*67*
10	*S+V*	*16*		*9*			*36*	*2*	*6*	*1*	*70*
11	*DW*			*9*			*36*	*6*	*4*	*1*	*56*
12	*TGV*					*42*	*6*	*2*	*2*	*1*	*53*
Total		*123*	*125*	*76*	*124*	*115*	*113*	*61*	*65*	*29*	*831*

List of milestones

The project has a duration of 36 months, which will be split into 5 milestones. These milestones represent a cycle of research, development as part of an overall architecture and evaluation through the scenarios. This is illustrated in Figure 16.

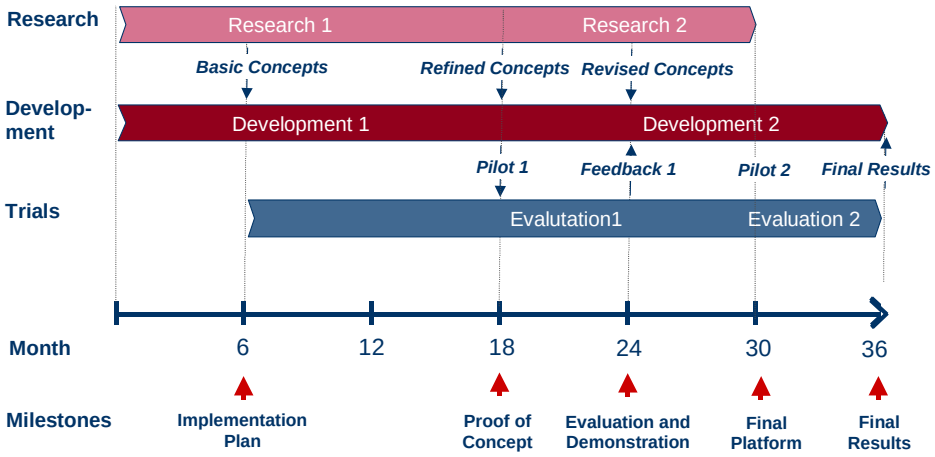


Figure 16: Project milestones

Table 7: Milestones

Milestone Number	Milestone name	Work package(s) involved	Expected date [24]	Means of verification[25]
1	***Implementation plan***	***1, 2, 3, 4, 5, 7, 8***	***Month 6***	***D1.1, D2.1, D3.2, D4.1, D5.1, D7.2, D8.1***
2	***Proof of concept***	***5, 7, 8***	***Month 18***	***D5.3, D7.3, D8.3***
3	***Evaluation and demonstration***	***5, 6, 7, 8***	***Month 24***	***D5.4, D6.2, D7.4, D8.5***
4	***Final platform***	***1-8***	***Month 30***	***D1.3 (M27), D2.6, D3.7, D4.6, D5.6, D6.4, D7.6, D8.7***
5	***Final results***	***1-8***	***Month 36***	***D1.4, D2.7, D3.8, D4.7, D5.7, D6.5, D7.7, D8.8***

[24] measured in months from the project start date (month 1).
[25] Show how both the participants and the Commission can check that the milestone has been attained. Refer to indicators if appropriate.

Section 2. Implementation

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 ConnectME, 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.

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.

Management structure

The management structure of ConnectME is illustrated in Figure 10. The **project coordinator** will have overall responsibility for the project. **Dr. Rüdiger Klein (Fraunhofer)** is an experienced researcher and project coordinator who provides all necessary capabilities for this role. 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.

ConnectME 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.

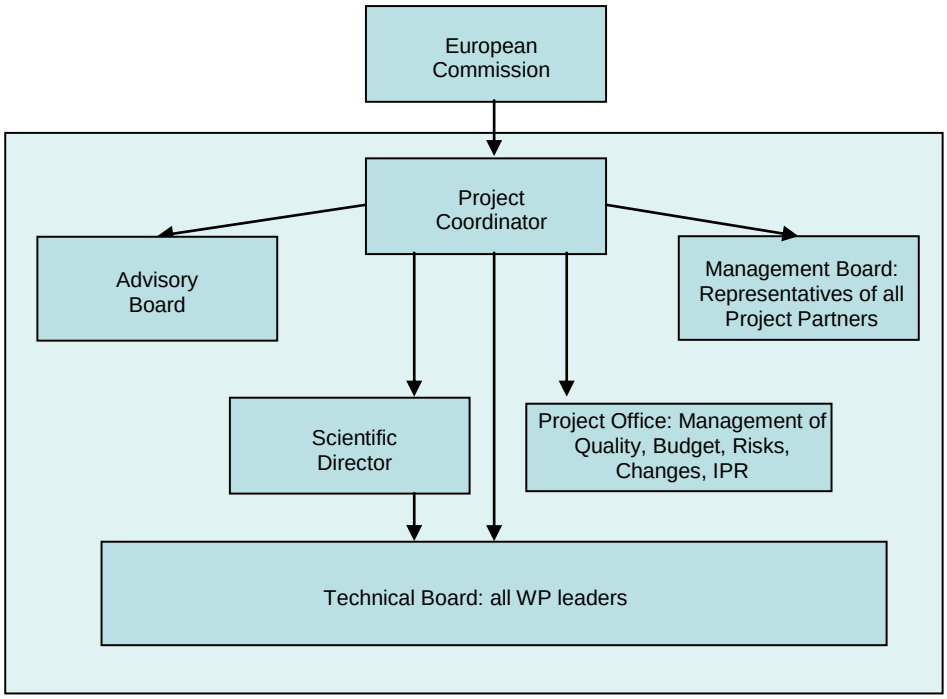


Figure 17: ConnectME Management Structure

Meetings and communication

Partners in ConnectME 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.

ConnectME 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 ConnectME 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. other 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.

Decision process

The decision making process in ConnectME 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.

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 ConnectME 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: Risks and contigency plan

Risk	Potential for risk	Impact on project	Contingency plan
Consortium – a partner can no longer provide the resources or skills foreseen in the workplan	Medium – personnel changes can not be avoided but all partners are very committed to the research themes of the project	High – work may no longer be done on time	The ConnectME consortium is prepared and agreed to imme-diately 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 ConnectME work	Low – while *similar* results may be launched, it is unlikely others will be able to replicate the expertise brought together in ConnectME to realise intelligent information services based on the dynamic interlinking of Web content	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 ConnectME 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 ConnectME
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 ConnectME 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, results may have to be produced outside 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 ConnectME 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, ConnectME may choose as an alternative to take other solutions which can provide at least part of the required results.	ConnectME has a consortium of partners with significant research work in the relevant areas and, if necessary, parts of the ConnectME 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 – ConnectME results fail to be taken up in the market	Low – while an innovative result does not always guarantee commercial success, ConnectME aims at achieving a multimedia delivery platform which is clearly a valuable goal for broadcasters and WebTV producers. The scenarios ensure that a public broadcaster and a cultural heritage archive in Europe will showcase the results.	Medium – while immediate exploitation of the results is the goal of ConnectME, it may be that the technology is still too immature at the end of the project for commercial uptake, or 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)	The exploitation task runs throughout the ConnectME project and is in a strong position to measure interest in the broadcasting and WebTV communities through our industrial partners. It is also able to quickly respond to changes in market outlook, as well as work strongly to influence that market itself with ConnectME dissemination and standardisation activities.

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.

Website and internal communication

As foreseen in the dissemination task, ConnectME 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 ConnectME 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 ConnectME 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 ConnectME 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 ConnectME 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 ConnectME 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 ConnectME but are highly dependent on regular usage by the consortium.

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 ConnectMe 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)

CERTH-ITI was founded in 1998 with its head office located in Thessaloniki, Greece. The most important related areas of R&D activities performed by CERTH-ITI include: semantic multimedia analysis, multimedia indexing and retrieval, multimedia and the Semantic Web, knowledge structures, languages and tools for multimedia, reasoning and personalisation for multimedia application, knowledge discovery for semantic-web applications, intelligent human computer interaction and intelligent agents, MPEG-7 and MPEG-21 standards. The Thessaloniki-based Information Processing Laboratory (IPL) of CERTH-ITI and its Athens-based Image, Video and Multimedia Systems Lab (IVML) of the National Technical University of Athens (NTUA), directed by professor Stefanos Kollias, will collaborate in the project. Together they have participated in more than 90 EC IST projects and 110 National projects and subcontracts. Over the last eight years, the CERTH-ITI research team has authored over 350 publications in scientific journals, 40 book chapters and over 520 presentations to international conferences in the multimedia and knowledge technologies field. For a complete list of research activities, R&D projects and publications, see http://www.iti.gr and http://image.ntua.gr. The two labs are coordinating the IST-FP7 Integrated Project "WeKnowIt" and are leading the "Media Intelligence" area. They have also participated in the IST-FP6 Integrated Project "aceMedia", the IST-FP6 NoE "Knowledge Web" and in COST 292 "Semantic Multimodal Analysis of Digital Media". They also have leading roles in the projects: "MESH: Multimedia Semantic Syndication for Enhanced News Services", "X-Media: Knowledge Sharing and Reuse Across Media", "K-Space: Knowledge Space of Semantic Inference for Automatic Annotation and Retrieval of Multimedia Content", "BOEMIE: Bootstrapping Ontology Evolution with Multimedia Information Extraction", "VIDI-Video: Interactive semantic video search with a large thesaurus of machine-learned audio-visual concepts" and others. Both laboratories are participating in standardisation activities including the W3C Multimedia Semantics Incubator Group, and conference organisation including CBMI 2009 and CIVR 2009.

Dr. Yiannis Kompatsiaris received the Diploma degree in electrical engineering and the Ph.D. degree in 3-D model based image sequence coding from Aristotle University of Thessaloniki (AUTH), Thessaloniki, Greece in 1996 and 2001, respectively. He is a Senior Researcher with ITI, Thessaloniki and currently he is leading the Multimedia Knowledge Group. His research interests include multimedia content processing, multimodal techniques, multimedia and the Semantic Web, multimedia ontologies, knowledge-based analysis, and context aware inference for semantic multimedia analysis, personalization and retrieval. He is the co-author of 6 book chapters, 19 papers in refereed journals and more than 60 papers in international conferences.

Dr. Yannis Avrithis received the Diploma degree in Electrical and Computer Engineering (ECE) from NTUA in 1993, the M.Sc. degree in Communications and Signal Processing from the Imperial College of Science, Technology and Medicine, University of London, in 1994, and the Ph.D. degree from NTUA in 2001. He is currently a senior researcher at the IVML of

NTUA. His research interests include image / video segmentation and interpretation, knowledge-assisted multimedia analysis, content-based and semantic indexing and retrieval and personalization. He has published 1 book, 13 articles in international journals, 5 book chapters, and 62 in conferences and workshops.

Dr. Phivos Mylonas received the Diploma degree in Electrical and Computer Engineering (ECE) from NTUA in 2001, the M.Sc. degree in Advanced Information Systems from the National & Kapodestrian University of Athens (UoA) in 2003, and the Ph.D. degree from NTUA in 2008. He is currently a researcher at the IVML of NTUA. His research interests include content-based information retrieval, visual context representation and analysis, knowledge-assisted multimedia analysis, issues related to multimedia personalization, user adaptation, user modelling and profiling. He has co-edited 5 books and 3 international journals, published 21 articles in international journals and books and 35 in international conferences and workshops.

Dr. Konstantinos Rapantzikos received the Diploma and the MS degree in Electronic and Computer Engineering from the Technical University of Crete (Greece) in 2000 and 2002 respectively and the PhD degree from NTUA in 2008. He is currently a researcher at the Image, Video and Multimedia Laboratory (IVML) in NTUA. His interests include modeling of visual saliency, salient/interest point detection in images/videos, biomedical imaging and motion estimation in compressed and uncompressed video. He has published 7 articles in international journals and books and 20 in proceedings of 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 KDD area (5FP projects Sol-Eu-Net and MiningMart), in medical informatics (4FP project MGT and DG SANCO project MedIEQ), in the TEL 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 PetaMedia, Knowledge Web, Ontoweb, KDnet, or EUNITE, and its members participate in several W3C working groups. The group also hosted top-class international conferences such as ECML (1997), PKDD (1999), EKAW (2006) and ISMIS (2009).

The key expertise with respect to ConnectME is related to *web mining and information extraction*, as well as to the issues of *multimedia-text analysis complementarity* and *multimedia ontology/metadata systems*, which will be exploited in WP1 and WP2 of ConnectME. 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.

CVs of key personnel

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, ISWC, ESWC, ASWC, ISMIS, SAMT, etc.); co-organiser of workshops held at ECML/PKDD, ISWC, ESWC and MIE. Member of the W3C OWL WG.

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. Recently he led experiments on applications of fuzzy rule reasoning for semantic region merging in multimedia analysis (within the K-Space project).

4. 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.

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.

5. 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).

6. 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.

7. **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.

8. 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 48 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 Seekda GmbH. He 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, including 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 bEU 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 several publications and has organized a number of workshops around related themes.

9. 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 ConnectME 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 Internet- 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).

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. Deutsche Welle

Deutsche Welle (DW) is Germany's international public service broadcaster, founded in 1953. In over 55 years, DW has gained its reputation as a trusted and reliable source of information by providing high-quality programmes and services in 30 languages to a worldwide audience. While, in the early days, Deutsche Welle provided its services primarily via analogue short wave radio, DW nowadays provides its audiences with up-to-date news and in-depth background analysis, offering truly multi-medial and multi-lingual programming. In addition to supplying "traditional" platforms and channels such as radio, television and online, DW is increasingly providing its services for emerging new devices (e.g. mobile phones, IPTV platforms, MP3 players etc), and respectively adapts the content to meet both devices capabilities and respective usage patterns. Deutsche Welle has been an active participant in European R&D projects since 1998. The primary aim is to use this involvement to stay at the forefront of developments in the media and IT sector in order to constantly improve the DW services and ways in which they are brought to the DW audience. ConnectMe is seen as an ideal project to aid DW in achieving just that, as it tackles some of the still unsolved challenges that result from digitisation and the emergence of new platforms.

Wilfried Runde worked as an Information specialist, researcher and TV-journalist for German broadcaster Westdeutscher Rundfunk (WDR) in Cologne, and for the ARD (Association of German Public Broadcasters) in the Brussels, New York and Washington bureaux. In 2001 Wilfried moved to Deutsche Welle. Ever since, he has managed and led various Research & Development projects that have addressed various aspects of the media business, ranging from interactive TV to workflow management in media organisations to the semantic web, to name just a few of the areas he has been involved in. Especially Wilfried's experiences gained in the MESH IP (www.mesh-ip.eu, completed on 28 Feb 2009) should be of great value to the work of the ConnectMe consortium. In ConnectMe, Wilfried will be primarily responsible for administrative issues and overall project co-ordination.

Jochen Spangenberg gained his very first active media experiences at the student-run Bath University radio station and during various placements while at University in Berlin. He then worked as a journalist and researcher for BBC Radio and TV, London. This was followed by five years in the (then) "New Media Economy" (1998-2003). In these rather formative years, Jochen held positions as Head of Media Consulting (including the management of R&D projects), Chief Operating Officer (in charge of all streaming media operations) and Editor-in-Chief (managing and being responsible for all sports productions for various platforms). Then, in 2003, Jochen joined Deutsche Welle where he has had a dual role ever since: leading DW's R&D projects together with Wilfried Runde, and working as a member of the DW strategy development team. In total, Jochen has been involved in around 20 EC co-funded R&D projects to date. In addition to his work for DW, Jochen regularly speaks at conferences about digital media production and delivery issues and is a Visiting Lecturer at various universities. Jochen's role in ConnectMe will focus on dissemination and exploitation activities.

Tim Koch has been involved in the running and management of R&D projects, many of them EC co-funded, since 2001. He joined Deutsche Welle in 2003 and, ever since, has been a member of the Innovation Projects team that handles and manages both EC co-funded research activities and other innovative projects at DW. Before he took up his work at DW, Tim was a task manager in the IST project iTV ("Designing, Authoring and Producing Enhanced Televised Content on the Internet and on TV" - one of the first projects dealing with interactive television and web services, completed in early 2002) and is currently responsible for part of DW's work in the ICT project CASAM (a project investigating computer-aided annotation of multimedia, www.casam-project.eu) and the eContentplus Programme project EUscreen (an Internet portal of various European TV stations and archives). In ConnectMe, Tim will be responsible for the co-ordination of DW's technical work and the validation activities.

The core team listed here will be supported by additional DW staff who will be involved in the technical work and validation activities. The individuals in charge of DW's ConnectMe work will ensure the smooth running of the project and do everything in their endeavour to make it a success.

12. **Thomson Grass Valley**

Thomson Grass Valley is a 100 % subsidiary of Thomson SA, the well-known professional electronics leader. It is a leader in the field of TV and radio terrestrial transmission, platforms for IPTV services, video-on-demand, Mobile TV, and Digital TV and radio broadcasting systems & equipment. TGV capitalises on advanced hardware, software and system capabilities and offers a wide variety of products and expertise in the fields of encoding, multiplexing, scrambling, conditional access management, streams syntactic analysis, SI / PSIP management, Electronic Program Guides, interface products, data broadcast, network supervision, streams servers, etc. To the best benefit of its customers, TGV plays the game of open standards: MPEG-2, MPEG-4, DVB, ATSC. TGV boasts first class references all over Europe, with customers as highly known as France Telecom, Deutsche Telekom, BBC, RBB, Canal +, Telefonica, ARD /ZDF, Mediakabel, Teledanmark, Eutelsat, SES Astra, Matra Grolier Network, etc.
TGV has been and still is a participant in key European research project, such as ADTT, FP4, FP5, FP6, FP7, Eureka and is a member of ISO MPEG, DVB and DAVIC. TGV therefore aims to play a key role in the multimedia head-end business and to efficiently support its customers' strategies in this migration towards digital. Skills and Know-How relevant to the head-end business are :

- MPEG encoding and decoding for SDTV or HDTV applications,
- Video repurposing to address mobile terminals,
- Multiplexing / demultiplexing / remultiplexing of digital transport streams,
- Scrambling / descrambling, conditional access EMM and ECM management,

- Real time and non-real time streams analysers and test generators,
- Global and local networks test and monitoring systems,
- Head-End configuration & monitoring software,
- Workflow & Asset Management software solution.

http://www.grassvalley.com/
http://www.thomson.net

Raoul Monnier is graduated from "L'Ecole Supérieure d'Electricité" of Paris (1983). From 1985 to 1987, he worked at Thomson CSF DTC in Cholet on software development for a radio surveillance system. In 1987, he joined Thomson CSF/ LER where he took responsibilities in theoretical studies, simulation and hardware design of a COFDM modem for digital TV broadcasting. From 1992 to the end of 1995, he worked on several RACE II projects where he was the representative of Thomson and project coordinator of studies and hardware developments. From 1998 to 2002, he managed the Front End and Signal Processing activity in Thomson R&D Rennes. He was responsible for the design and development of the Front End part of all the STBs developed by Thomson for the European market (Canal+, TPS, Golden Channels, ...). From 2002 to 2006, he managed several Advanced Development projects, including Thomson's developments for SATMODE (European Space Agency project). Since December 2006, he has been European Collaborative Project Manager and has been managing for Thomson Grass Valley several FP6/FP7 projects, such as Imosan and NoTube and French national collaborative projects.

Christophe Berthelot is an audio/video Architect Manager in the NTA Department of Thomson Grass Valley. He received a degree in Computer Science and Signal Processing from the INSA Engineering school of Rennes (France) in 1994. Currently, he is managing the definition of a new audio/video architecture platform for audio/video processing in the compressed and uncompressed domain and publishing. He has played the leading role in developing splicing, transrating and logo insertion algorithms for MPEG-2 and AVC video. He has a strong knowledge in audio/video/system compression and distribution standards and advanced platform architecture.

Consortium as a whole

The success of ConnectME 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 SME which stands to benefit from ConnectME market opportunities and large traditional media players who will benefit from the innovation of ConnectME technology to strengthen their respective market position (content provider, broadcaster, network services).

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 ConnectME, 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 ConnectME. 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 ConnectME. **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 ConnectME 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 ConnectME user interface. For implementation of the ConnectME front end on an interactive video player CWI can take the AMBULANT video player[26] as a basis.

The POLETI department of the **University of Mons (BE)** has acquired an expertise in computational attention which provides a non-intrusive way of behaviour and reaction analysis. This expertise is needed for workpackage 4 and more precisely for task T4.2. The proposed models focus on efficiency and real time implementation in ecological situations more than on artificially generated data. Several models of both bottom-up and top-down attention were implemented on still images, but also on 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 have been set up at POLETI/UMONS in the framework of multimedia data as music, images, videos or laughter signal. The multimedia data retrieval techniques are needed in workpackage 1 in coordination with the other partners.

[26] http://www.ambulantplayer.org/

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 ConnectME, 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 ConnectME 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 ConnectME 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 ConnectME.

For development of the back-end infrastructure, including metadata repositories and service endpoints, and to support the broadcasters in integrating ConnectME 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.

It is of vital importance in ConnectME 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 ConnectME, the public broadcaster **Deutsche Welle (DE)** will provide television programming as well as facilities for user trials and potential for technology exploitation to the project.

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

Finally, **Thomson Grass Valley (FR)** will provide the platform hosting the software modules developed in the project and allowing to stream the IPTV programs toward the end users.

Resources to be committed

ConnectME work will be done with a total of 831 person months over 36 months, representing ca 23 persons' effort per month with 12 partners asking for funding. The total funding requested is approximately 6.5 Million €.

ConnectME is an Integrated Project. It is directed towards innovations in one of the central fields of Intelligent Information Management where expertise from different domains has to be integrated: that of handling interlinking of meaning between different (non-textual) media. We have partners from industry including a SME, a public broadcasting organization, a public cultural archive and a multinational in the electronics field, as well as 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, UMON and EURO contribute their methodological research and experiences here in joint efforts. WPs 1, 2 and 4 have a comparable workload, as significant research effort is needed to extend the theoretical research the partners are experts in into new scalable methodologies and systems for ConnectME. In WP3 (ConnectME interface and presentation engine) the workload is slightly less as one partner (CWI) will solely contribute and extend existing resources and expertise in interface design (as opposed to theoretical research based on cross-partner collaboration), with the scenarios as well as vision of Connected Media providing the requirements.

WP5 is dedicated to the creation of the ConnectME platform. This is typically a bigger effort including large amounts of software engineering and integration. Our three partners Condat, TGV, and CWI 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.

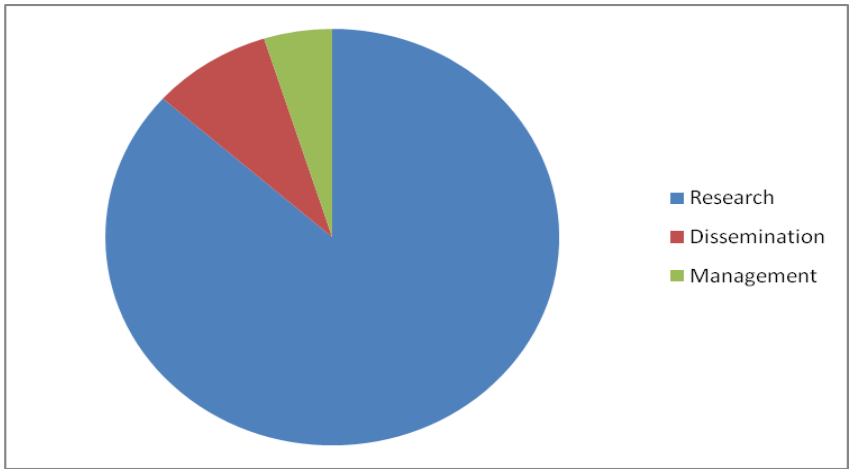
WP6 is dedicated to the creation and evaluation of our use cases. Both use case partners DW and S+V will use their 36 PM to prepare the use cases raw material, to define the requirements, specify the scenarios, contribute to their implementation on the ConnectME platform, and evaluate the fulfilment of their initial requirements. They will be supported by STI which will take the lead for this WP6. This guarantees that both can contribute to the overall approach of ConnectME 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 61 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 EURECOM 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 65 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 implementation SME and use case partners will actively support these activities. Additionally, Fraunhofer as IP coordinator and STI as research network contribute to exploitation.

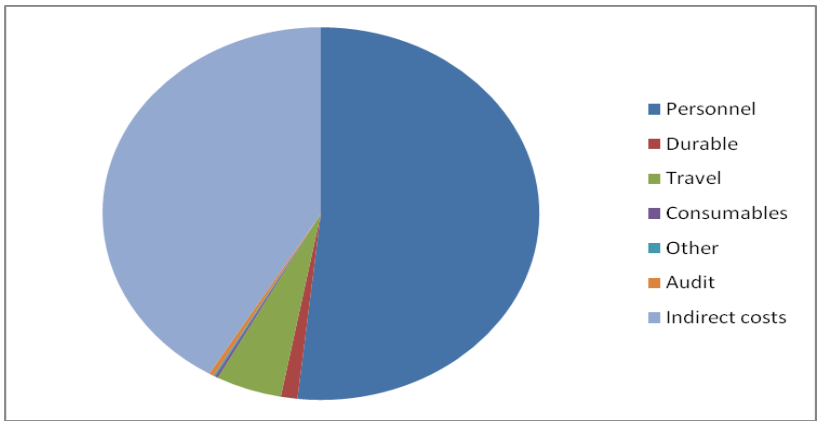
The management WP9 is deliberately lean. The main effort in this area will be provided by Fraunhofer as IP coordinator. All other partners will get a small amount of work load for their management activities. They will be tightly integrated with all other activities in the IP.

The breakdown between research, dissemination and management in the budget is shown below.



Research is the major part of ConnectME (90%) however we give a significant focus to dissemination of results (6%). The remainder of the budget (4%) is allocated to appropriate and effective management.

This budget is split between different cost items:



It can be seen that personnel makes up most of the requested budget (52%), the other larger cost item being the overhead (indirect costs) (41%). These ensure that sufficient personnel are committed to carrying out the workplan of the project and have access to all necessary infrastructure at their institutions.

From the other cost items, travel is budgetted to allow for attendance of project meetings as well as dissemination activities (5%). Other cost items represent minor costs (audit certificates, or consumable equipment). We mention in particular durables (1.2%) as we budget for the use of cloud computing services for processing power and Web mining (e.g. Amazon EC2, 80legs) at a larger scale than would be possible even on institution's own networks. This allows us to achieve scale of billions, at a lower cost than purchasing high end computing equipment.

Section 3. Impact

One central result of ConnectME 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 ConnectME: in society, and in business.

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 ConnectME. 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 the rich media Web. This new paradigm merges the popularity and accessibility of video and television as a means of communication with the benefits of the Web as an immense source of information and content. 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 ConnectME 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 ConnectME services. The consortium has prepared itself well for all these challenges, as outlined beforehand and described in detail in this document.

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 video 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. ConnectME 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.

Business: strengthening Web video and TV in Europe

ConnectME can provide the technologies to place Europe at the innovative forefront of interactive Web and IP TV/video service creation and provision. The distribution of Web-based TV and video today is dominated by US distribution platforms. Leading platform is YouTube with almost three quarters of Web video offerings (see Figure 17 below):

Top Online Brands ranked by Video Streams for September 2009 (U.S.)			
RANK	Video Brand	Total Streams (000)	Unique Viewers (000)
1	YouTube	6,688,367	106,180
2	Hulu	437,407	13,519
3	Yahoo!	228,494	30,084
4	MSN/WindowsLive/Bing	180,104	18,109
5	Fox Interactive Media	139,634	14,342
6	Nickelodeon Kids and Family Network	127,654	5,303
7	Turner Sports and Entertainment Digital Network	123,665	6,062
8	MTV Networks Music	116,839	9,647
9	ESPN Digital Network	114,652	9,299
10	Facebook	110,418	23,161
			Source: The Nielsen Company

Figure 17: Top Online brands ranked by video streams for September 2009 (USA)

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 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[27].

The convergence of audio-visual content and other Web content can 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 technology a competitive advantage in a growth market and support the digital society by enabling new forms of information access to citizens.

We believe that the converged services that can be made possible over the ConnectME framework can hold the key to establishing European leadership in Web-based TV and video technology and to unlocking the "goldmine" of new compelling rich media Web applications. 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[28]. (cf. Exploitation section below)

Why Europe ?

The work of ConnectME 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.

Why will we succeed?

With regard to other research activities, the research community relevant to ConnectME 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 ConnectME. 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 ConnectME 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 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 ConnectME: 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 ConnectME, 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 ConnectME strong co-operation opportunities. In particular, standardisation activities will be carefully monitored and to the best of the partner's

[27] 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
[28] http://www.tvover.net/2009/02/26/IPTV+Leads+The+Telco+Transformation+In+2009.aspx

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 interactive TV and video, and related sectors of relevance to the ConnectME work, and ensure ConnectME's uniqueness in what will be a rapidly growing market. As partners come from the broadcasting, content provision, network operation, and 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 ConnectME 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 ConnectME objective (see also the state-of-the-art section for an indication of what can be expected in the next 1-3 years)

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 the Web (on TV, mobile, other devices). As a result of ConnectME, 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 Linked Media layer at Web scale. This can be extended by user and context-personalised aggregation of relevant associated data, e.g. through user eye tracking, the concept-related content is further refined to emphasize the objects users pay more attention to. 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, where there is potential here to extend the framework to the ubiquitous Web, which may include mobile devices, home devices and even virtual reality-like environments. ConnectME 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 ConnectME is expected to be a society in which video streams and Web converge in any device, and are 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 viewing of audiovisual content such like television – can be an educational and stimulating activity rather than breeding a new form of "online couch potato".

The establishment of the ConnectME 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, leading to a paradigm shift in audiovisual content which brings in the interactivity that made the Web of text and images so popular, and avoiding the plague of passivity which has inflicted many viewers of video, used to the television setting.

For the citizen, ConnectME 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, ConnectME can mean market leadership in a promising growth sector (Web-based video distribution and consumption). For European society, ConnectME can be the basis for ubiquitous and intuitive access to relevant information through Web-converged TV broadcasting.

The long term outlook for ConnectME

ConnectME will create a platform for hypervideo, which can be to the future Rich Media Web, just as HTML was necessary for the success of hypertext on the first generation World Wide Web.

ConnectME services represent a new interaction paradigm with audiovisual content which can be seen as the fulfilment of hypervideo[29]. It will be able to provide the platform for interactive services on future content networks, beyond our initial application of Web-based and broadcaster-sourced content to PC, onto IPTV, home networks and mobile devices. Combined with future technologies like 3D internet and virtual realities, *the ConnectME platform will deliver hypervideo content and applications to users of the future Rich Media Web much as HTML formed the core technology of the first generation World Wide Web experience.*

Dissemination and exploitation of project results, and management of intellectual property

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 ConnectME 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.

[29] 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

- Establishment of close links with other FP7 projects active in the same areas as ConnectME 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 ConnectME 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 ConnectME 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 ConnectME. 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 ConnectME 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 ConnectME work will be disseminated, include:

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

- GMF - The Deutsche Welle 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
- CIVR - ACM International Conference on Image and Video Retrieval
- ICIP - IEEE International Conference on Image Processing
- ICME - IEEE International Conference on Multimedia & Expo
- EUSIPCO - European Signal Processing Conference
- ICSC - IEEE International Conference on Semantic Computing
- Museums and the Web
- International Federation of Television Archives (FIAT-IFTA) annual conference
- International Broadcasting Convention (IBC)
- European Conference on Digital Libraries
- Relevant events organised by EBU, Europeana, Open Video Alliance

It is planned that ConnectME 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 ConnectME 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
- 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 ConnectME are involved in relevant FP7 projects indicating co-operation and cross-project dissemination opportunities. Besides their partnership in ConnectME, consortium members are also involved in following related R&D projects:

- X-Media – Large Scale Knowledge Sharing and Reuse Across Media (2010)
- PAPYRUS. Cultural and historical digital libraries dynamically mined from news archives." (2010)
- CASAM: Computer-Aided Semantic Annotation of Multimedia (2011)
- NoTube: Networks and Ontologies for the Transformation and Unification of Broadcasting and the Internet (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. ConnectME'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 ConnectME will have substantial impact on further standardization efforts. ConnectME partners are uniquely positioned to contribute due to their past and existing activities and contributions in the area. For example: CERTH-ITI actively participates in standardisation 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".

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 ConnectME. 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 ConnectME. 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, Web video and TV portals.

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.

Exploitation by partners

It is planned to pursue a two-way exploitation strategy in ConnectME:

- 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 ConnectME 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 ConnectME 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 ConnectME 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 ConnectME 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:

Market trends

Web-based TV and video is a growth market globally. For example in October 2009 NielsenWire reported that "*Year-over-year, unique viewers, total streams, streams per viewer*" were increasing, "*…led by 25% percent growth in total streams and time per viewer*"[30]. In August 2009, a record 25 billion videos were watched online. On October 9, 2009 YouTube announced that it was now serving 1 billion videos per day[31].

In parallel to the developments of online video, the global IPTV subscriber base is expected to grow from 20 million in 2008 to almost 90 million in 2012. Already, Europe is being heralded as the potential market leader in IPTV[32] 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.

In general, the overall media consumption is increasing. The growth is partly resulting from media-multitasking, i.e. simultaneous consumption of media on three different screens - classical TV, online and mobile TV[33]. One emerging trend in this context are the so called "Telecommunities", which comprise people who simultzaneously watch TV and chat real time[34]. For example, according to NielsenWire, 11% of people who watched the Academy Awards were logged onto the Internet (see footnote 32). This are indicators that consumers are already transforming video consumption from a lean-back to a lean-forward medium, and show affinity to interactive use as well as towards extending video content with additional information and activities.

In order to gain the leading role in the Web video and TV market, European platforms and operators must begin to leverage the characteristics of the new platform with interactivity to produce a differentiated offering, redefining the experience of video consumption. At the same time innovative technologies as ConnectME would enable European companies to create niche interactive web TV markets not available yet.

[30] Time Spend Viewing Video online Up 25% per Viewer, Nielsen wire, 13.10.2009 http://blog.nielsen.com/nielsenwire/online_mobile/time-spent-viewing-video-online-up-25-per-viewer/
[31] http://youtube-global.blogspot.com/2009/10/y000000000utube.html
[32] Europe to dominate IPTV growth, Light Reading, 5.3.2007 http://www.lightreading.com/document.asp?doc_id=118668
[33] see for example the report „Three Screen Report: Media Consumption and Multi-tasking cantinue to increase Across TV, Internet, and Mobile". NilsenWire, 2.09.2009 http://blog.nielsen.com/nielsenwire/online_mobile/three-screen-report-media-consumption-and-multi-tasking-continue-to-increase/
[34] Media is on Demand - But Content is still king. NielsenWire, 07.04.2009 http://en-us.nielsen.com/main/insights/consumer_insight/april_2009/media_is_on_demand

According to a major market research company, the main options how to gain competitive advantage and market share in the IPTV sector are[35]:

- 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."

Exploitation Strategies of ConnectME Partners

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

ConnectME, will deliver new technology which will allow web TV, video and IPTV broadcasts to be radically different from traditional TV broadcast and from current video and Web TV offerings through the integration of Web technology. Telcos offering services based around ConnectME 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 ConnectME to enhance their broadcasts in dynamic, intelligent and customizable fashion. Through the ConnectME-enabled analysis and annotation of broadcast material, the set of ConnectME technologies can be used to create added value services without prohibitive manual preparation. Media providers can increase market share by offering their content over ConnectME 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 ConnectME, interaction with video content will reach maturity as usage will be based on the content of the visual material – 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 ConnectME 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. The Multimedia Knowledge Laboratory of CERTH-ITI is active in providing research services and results to the local and European industry through direct research contracts and licensing agreements.

[35] http://www.abiresearch.com/press/1127/IPTV+Operators+Searching+for+Competitive+Differentiation

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 ConnectME 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 ConnectME 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 ConnectME 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.

Primarily, **Deutsche Welle** will exploit the ConnectME 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 ConnectME technology in its workflow, any by offering it to the content production staff, Deutsche Welle 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.

In addition to the internal exploitation plans outlined above, DW 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, Deutsche Welle 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, DW can well imagine establishing closer partnerships with other consortium members on a separate legal level (i.e. form a new legal entity - something DW is entitled to by its statutes) in order to deal with the full-scale exploitation of project achievements.

ConnectMe 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: ConnectME exploitation opportunities

Potential Exploitation Partner	ConnectME 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	ConnectME platform for annotation & content aggregation	Further development, documentation, licensing, monetarization opportunities (server-side advertising)
Software/Services for TV/Video playback	ConnectME interactive video player	Further development, documentation, licensing, monetarization opportunities (client-side advertising)
The European citizen	ConnectME services	Public promotion (e.g. online showcases), user trials, private service offers

Possible product niche for ConnectME

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 Web-based interactive video 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 ConnectME components to key market players.

Therefore, ConnectME should not only be exploited as a PC-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, linear TV is at present ported to mobile devices. Most of the European countries have introduced linear mobile TV. TVs are converging with Internet, with or without the separate Set Top Box, which are increasingly supporting more interactivity. In the future, we fully expect mobile video and IPTV to have a similar development path as Web video and become an early adopter market of these technologies. In particular, as early experiences with mobile TV and IPTV show that more sophisticated and personalized services that go beyond linear TV are required. No services comparable to ConnectME 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.

Knowledge Management

Knowledge management is very important in an innovative project such as ConnectME. 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 ConnectME to advance their own research and development, have a basis for protecting the knowledge gained for both future dissemination and exploitation.

While ConnectME 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 ConnectME platform will carry an open source license, so that further development by commercial organisations is supported. We will allow the software partner CONDAT to enhance their existing solutions with the main features developed in the project. The exploitation partners Deutsche Welle, CONDAT and Sound & Vision will also continue to have access to the platforms for ConnectME services developed in the project.

The consortium agreement will be the basis in ConnectME 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

None. Eye tracking technology will be used in user trials with participant's consent.

Table 10: ETHICAL ISSUES

	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. aceMedia (FP6-IP) http://www.acemedia.org/aceMedia
2. X-Media (FP6-IP) http://www.x-media-project.org
3. Muscle (FP6-NoE) http://www.muscle-noe.org
4. MESH (FP6-IP) http://www.mesh-ip.eu/?Page=Project
5. P. Maragos, "Cross-Modal Integration for Performance Improving in Multimedia: State of the Art Report" MUSCLE NoE Deliverable D.6.1, Sept. 2004.
6. Ph. Mylonas, E. Spyrou, Y. Avrithis and S. Kollias, Using Visual Context and Region Semantics for High-Level Concept Detection, IEEE Transactions on Multimedia, Volume 11, Issue 2, pp. 229-243, February 2009
7. Th. Athanasiadis, Ph. Mylonas, Y. Avrithis, S. Kollias, "Semantic Image Segmentation and Object Labeling", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 17, Issue 3, March 2007.
8. M. Boutell, "Exploiting Context for Semantic Scene Classification", Technical Report 894 (Ph.D. Thesis), University of Rochester, 2006.
9. Ph. Mylonas, D. Vallet, P. Castells, M. Fernandez and Y. Avrithis, Personalized information retrieval based on context and ontological knowledge, Knowledge Engineering Review, Cambridge University Press, Volume 23, Issue 1, pp. 73-100, March 2008.
10. http://www.riya.com/
11. http://www.youtube.com
12. http://www.asterpix.com/
13. http://www.videoclix.tv/
14. http://www.klickable.tv/
15. http://www.media.mit.edu/hypersoap/
16. http://www.overlay.tv/
17. http://www.blinkx.com
18. http://hadoop.apache.org/
19. http://www.archive.org/details/movies
20. 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.
21. Oomen, J. and V. Tzouvaras (2007), Providing Access to European Television Heritage. In Ariadne. Issue 53 http://www.ariadne.ac.uk/issue53/ooman-tzouvaras/
22. 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.
23. Dalton, M. S. and Charnigo, L. (2004). Historians and their information sources. College & research libraries 65:400-425.
24. Wright, Richard (2008) Preservation of Digital Audiovisual Content: briefing paper. Accesed online at: www.digitalpreservationeurope.eu/publications/briefs/audiovisual_v3.pdf
25. 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.
26. K-Space NoE, http://www.k-space.eu/
27. Boemie, http://www.boemie.org/
28. 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

29. 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.
30. 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.
31. 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.
32. 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.
33. 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.
34. 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.
35. 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
36. 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.
37. 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.
38. 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.
39. 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.
40. 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.
41. 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.
42. 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.
43. 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
44. http://www.metadataworkinggroup.com/
45. http://www.w3.org/2008/WebVideo/Annotations/
46. L. Hardman. Modelling and Authoring Hypermedia Documents. PhD thesis, CWI, Amsterdam, 1998.
47. 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.

48. 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.
49. 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.
50. Synchronized Multimedia Integration Language (SMIL 3.0). W3C Recommendation, 01 December 2008, http://www.w3.org/TR/SMIL/
51. 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 http://youtube.com/watch?v=UxnopxbOdic
52. 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.
53. 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.
54. 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
55. Labský, Martin, Svátek, Vojtěch, Nekvasil, Marek, Rak, Dušan. The Ex Project: Web Information Extraction Using Extraction Ontologies [online]. In: Berendt, Bettina, Mladenič, Dunja, Gemmis, Marco de, Semeraro, Giovanni, Spiliopoulou, Myra, Stumme, Gerd, Svátek, Vojtěch, Železný, Filip. Knowledge Discovery Enhanced with Semantic and Social Information. Berlin : Springer-Verlag, 2009, s. 71–88. 148 s. Studies in Computational Intelligence 220/2009. ISBN 978-3-642-01890-9. ISSN 1860-949X. URL: http://springerlink.com/content/j28100148190/.
56. Kliegr, Tomáš, Chandramouli, Kriishna., Nemrava, Jan, Svátek, Vojtěch, Izquierdo, E. Combining Image Captions and Visual Analysis for Image Concept Classification. Las Vegas 24.08.2008 – 27.08.2008. In: MDM 2008 Workshop – 9th Imternational Workshop on Multimedia Data Mining. [online] Las Vegas : ACM, 2008, s. 8–17. ISBN 978-1-60558-261-0. URL: http://portal.acm.org/citation.cfm?id=1509212.1509214&coll=ACM&dl=ACM&CFID=2833786&CFTOKEN=37372318
57. Ossenbruggen, J. R., Cornelissen, F. J., Geurts, J. P., Rutledge, L. W., and Hardman, L. 2000 *Cuypers: a Semi-Automatic Hypermedia Generation System*. Technical Report. UMI Order Number: INS-R0025., CWI (Centre for Mathematics and Computer Science).
58. Lindley, C. A., Davis, J. R., Nack, F., and Rutledge, L. W. 2001 *The Application of Rhetorical Structure Theory to Interactive News Program Generation from Digital Archives*. Technical Report. UMI Order Number: INS-R0101., CWI (Centre for Mathematics and Computer Science).
59. Hildebrand, M., van Ossenbruggen, J., Hardman, L., Jacobs, G., Supporting subject matter annotation using heterogeneous thesauri, a user study in Web data reuse. International Journal of Human-Computer Studies (IJHCS), 67(10):887–902, 2009.
60. Schreiber, G., et al., Semantic annotation and search of cultural-heritage collections: The MultimediaN E-Culture, Demonstrator. Journal of Web Semantics (JWS), 6(4):243–249, 2008.
61. Bordegoni, M.; Faconti, G.; Maybury, M.T.; Rist, T.; Ruggieri, S.; Trahanias, P.; Wilson, M., "A Standard Reference Model for Intelligent Multimedia Presentation Systems", Computer Standards & Interfaces, 18:477-496, 1997.

Appendix I: Full list of projects compared to ConnectME

Project description	Similarities and differences with ConnectME
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/	ConnectME will build on the results from this NoE, since four partners (CERTH, CWI, Eurecom and UEP) already collaborated here. ConnectME has a wider goal than just the focus on media analysis. K-Space did not have specific use cases while ConnectME will further develop analysis and annotation technologies tailored for the needs of Connected Media Experiences and evaluated in two 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. ConnectME will tackle cross-media integration and interaction using near real-time methods that do not rely 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 ConnectME it employs a quite different application domain (automotive industry) and does not address further challenges of search and retrieval of media in open large scale networks like the Web.
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;	ConnectME 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 two

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	specific scenarios. In particular, ConnectME 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 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	MESH is focused on a single application domain (i.e. news) and on the spatial segmentation of news media, while ConnectME develops a platform for generic multimedia services, implements two different scenarios and requires a finer grained spatio-temporal segmentation of different types of media.
INTERMEDIA — Interactive Media with Personal Networked Devices	The project focuses on the whole range of personal adaptations for multi-media sources onto different types of wearable devices for moving users. ConnectME 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 — Staging of Media Events	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 ConnectME.

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 ConnectME in the sense that both projects aim at the automation of the labeling process and interrelating media from multiple sources. ConnectME will deploy the speech recognition toolkit developed partly in this project that is currently being used within Sound and Vision for audio analysis. ConnectME 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 ConnectME.
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. ConnectME will go beyond the boundaries of the published content and link to Web-based media available from other sources. All interaction points in NM2 are pre-specified, whereas in ConnectME, 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 ConnectME focuses on a semi-automatic approach using intelligent approaches and re-using a Web-size metadata layer (Open Linked Data) as well as extending it with multimedia annotation (*Linked Media*). 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 for the long-term preservation of digital media	Related to ConnectME, PrestoPRIME will create a metadata conversion and deployment

objects, and find ways to increase access by integrating the media archive with European on-line digital libraries in a digital preservation framework. 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.	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 ConnectME (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 ConnectME, RUSHES does not tackle the lack of integration between audio-visual material and other 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 open, large scale Web dimension of ConnectME, which will support the linking between Web-based 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 ConnectME services.
SEMEDIA — Search Environments for Media	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. In addition, ConnectME will start much later and can therefore uptake the most recent research results regarding video analysis and similarity-based retrieval.
CHORUS — CHORUS is a Coordination Action which aims at creating the conditions of mutual information and cross fertilisation between the projects that will run under	Partners of CHORUS (CERTH) participate in ConnectME and therefore the project will benefit from the useful experience and the results of CHORUS. More specifically,

Strategic objective 2.6.3 (Advanced search technologies for digital audio-visual content) and beyond the IST initiative. http://www.ist-chorus.org/	ConnectME will reuse the state of the art reports on audio-visual search technologies produced by CHORUS 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 project cluster will be set up when necessary in order to collaborate and re-use existing technology.
IM3I — Immersive Multimedia Interfaces. http://imthreei.hku.nl/index.html The IM3I project develops interactive and innovative applications and services that support users in developing creative content to interact smoothly with multimedia data using advanced mixed multimedia search and retrieval tools and 3D results presentation and navigation interfaces.	The IM3I project shares similar goals with ConnectME in terms of understanding user needs for interacting with annotated multimedia, but focusses more on the automatic annotation and indexing of multimedia assets. ConnectME will go beyond the annotation and indexing processes and provide content-based linking mechanisms that provide users access to external networked media sources.
CASAM — Computer-Aided Semantic Annotation of Multimedia	ConnectME 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 ConnectME 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 ConnectME's goal of content aggregration and presentation. While there may be some co-operation possibilities in metadata-driven processing of digital media, ConnectME 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 is to increase the feeling of "connectedness" with other people or the perception of "involvement" in a particular shared experience. http://www.inem4u.eu/	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 ConnectME, where the intention is to provide multi-device access to content. Within ConnectME more emphasis is given to identifying connection points within 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 ConnectME through sharing research both in the broadcaster annotation tool and end user interfaces.
MYMEDIA — Dynamic personalisation of Multimedia	MYMEDIA emphasizes on the personalization of content consumption and recommendation of relevant content to the user. ConnectME 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 — P2P Tagged Media Four partners from the Netherlands, Switzerland, UK and Germany form the core of the PetaMedia NoE, each representing a national network. The goal is to bring together the research of these networks in the area of multimedia content analysis and social peer-to-peer networks and eventually to establish a European virtual centre of excellence. The research 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 combining these.	The Call 2 PetaMedia NoE shares with ConnectME the effort in exploiting complementary textual and social-network data for indexing multimedia. However, while PetaMedia primarily deals with back-end algorithms for media indexing, ConnectME pays significant attention to the user-facing front end. Furthermore, PetaMedia, as an NoE, aims at integrating existing research at different workplaces, while ConnectME, 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. One of the aims of the Petamedia NoE is to stimulate research connections, which has led to two partners (CWI and Eurecom) becoming affiliated partners.
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 ConnectME'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 were addressed in a number of sub-projects: Learning Features, Multimodal Interaction, Ambient Multimedia Databases, Semantic Multimedia Access, Professional's Dashboard,	Many of the results from MultimediaN are relevant to ConnectME. In particular, ConnectME will build directly on results on user interaction with semantically annotated cultural heritage repositories in the E-Culture project. Here information exploration tasks were investigated, in particular how image

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.	assets relate in different ways to concepts that are of interest to the user. ConnectME will go beyond these tasks to investigate how users can navigate through linked video assets. ConnectME 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. ConnectME is more oriented toward the end user (content personalization) with the main goal of developing new experiences of Web-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.
VidiVideo - The Vidi-Video project aims to integrate and develop state of the art components from machine learning, audio event detection, video processing, interaction and visualization into a fully implemented audio-visual search engine. The highlight of the project is the development of a large number of video detectors (in the order of 1000), each corresponding to one concept, thus allowing the association of each video shot with one or more of the approximately 1000 concepts. This will facilitate the concept-based retrieval of video in large collections. http://www.vidi-video.it/	ConnectMe partner CERTH is also a member of the Vidi-Video consortium and actively participates to it, gaining valuable experience in video detector development. As a result of this, ConnectMe will benefit significantly from the developments in Vidi-Video, since the large number of state-of-the-art detectors developed in the latter can serve as an excellent starting point for ConnectMe. However, at the same time there is significant ground to be covered before these detectors can be used towards the goals of ConnectMe. This is mostly due to a) real-time processing being a low priority in Vidi-Video, since the emphasis is on documenting the feasibility and usability of developing a large number of detectors, rather than on solving the problems related to their time-efficiency (i.e. real-time processing requirements) and their applicability to extremely large datasets; b) the detectors of Vidi-Video operating only at the shot level, whereas in ConnectMe the annotation of video not only at shot level but also at the object level is among the major requirements.