Small or medium-scale focused research project (STREP) proposal
ICT Call 9
FP7-ICT-2011-9
Content-based Access for Cinematographic Heritage Exploitation

CACHE-22 *

Small or medium scale focused research project (STREP)
Date of preparation: 14 april 2012
Version number *(optional)*: 00

Work programme topic addressed
(if more than one, indicate their order of importance to the project. The main (first) objective <u>must</u> be one included in this call)
Challenge 8: ICT for Learning and Access to Cultural Resources
ICT 2011.8.2 ICT for access to cultural resources

Name of the coordinating person: Vincenzo Lombardo
e-mail: vincenzo.lombardo@unito.it
fax: 011-751603

Participant no. *	Participant organisation name	Part. short name	Country
1 (Coordinator)	Università di Torino	UniTO	Italy
2	Centrum Wiskunde & Informatica	CWI	The Netherlands
3	Informatics & Telematic Institute	CERTH	Greece
4	IN2	IN2	UK
5	Athens Technology Center	ATC	Greece
6	Nederlands Instituut voor Beeld en Geluid	NISV	The Netherlands
7	Cinecittà Luce	LUCE	Italy

Please use the same participant numbering as that used in Proposal submission forms A2

Proposal abstract
(copied from Part A, if not in English include an English translation)

CACHE-22 will generate filmicly meaningful annotations that will enrich cinematographic heritage repositories. These will provide the basis for a variety of services, such as narrative-based search tools, the automatic creation of content-based links within and among films, and the creation of links with the cloud of open linked data.

Examples of users and services that the project will enable are:

- Scholars, museum visitors, teachers and students will be able to search the enriched film repositories on the basis of narrative functions, thus improving the access to European film heritage and its use for teaching purposes.
- Professional search services will be able to provide narrative-based access to film stocks.
- Cultural institutions will be able to construct interactive exhibitions of cinematographic heritage, allowing linking among aspects such as narrative role of a character, cinematographic technique or plot development.
- Tourist offices of cities featured in films will exploit the enriched repositories to provide cinema-based explorations of their cities.
- Film writing students will be able to actively learn the art of dramatic construction by (re-) using film fragments as examples of specific concepts, e.g. dramatic situations.

The basis of the project is that meaningful annotations of cinematic material will enable a wealth of applications. The catch is that
- without annotations, no-one can find the media to re-purpose it;
- without re-purposing tools, no-one will annotate the media.

CACHE-22 will enable automatic annotation chain and filmic concept enrichment to re-use existing cinematographic repositories.

Table of contents

Section 1: Scientific and/or technical quality, relevant to the topics addressed by the call
(Recommended length for the whole of Section 1 – twenty pages not including the Gantt chart, Pert diagram or tables 1.3a-e)

1.1   Concept and objectives

Explain the concept of your project. What are the main ideas that led you to propose this work? Describe in detail the S&T objectives. Show how they relate to the topics addressed by the call, which you should explicitly identify. The objectives should be those achievable within the project, not through subsequent development. They should be stated in a measurable and verifiable form, including through the milestones that will be indicated under section 1.3 below.

Film clips are search items on the web, objects of study for cinema students and academics, media assets in the production pipeline of audiovisual companies. Given the multiple roles of a film clip (respectively, as a web resource, as the manifestation of some narratological and filmic theory, and as a media asset in production), the annotation of narrative features in multimedia data is useful for several reasons:

- Access to cultural heritage archives. Narrative–based access can support institutions in providing services that are of interest to the general public and scholars, in a way that overcomes the limitations of the tag-based annotation and search and makes the role of the film archives more active in the society.
- Analysis of narrative media objects. The analysis of narrative media objects is part of the training in film and media studies. Annotated narratives can support this training, especially through dedicated IT tools for the methodology of analysis.
- Media production. Traditional media production undergoes several complex tasks, handled by different professionals (writer, director, cinematographer, editor, actors, ...). The annotation of the narrative features can be started during the authoring process and propagated through the production phases to make all the team aware of what is going on (common tools used for this goal are storyboards and screenplay breakdowns, that are not machine readable). The software tools currently employed to assist these tasks do not generated interoperable, shared representation of the story content.
- Search/Retrieval. The annotation of narrative features is important in the search/retrieval of individual media assets, that are component parts of some larger object. Again, this is relevant not only for the fruition of the filmic heritage, but also for media production itself (for example, for the editor to retrieve some footage shot), but also for content aggregators (for example, in the compilation of programme schedules or trailers).

Narrative media introduce real and fictional worlds by structuring events and characters into stories. As a consequence of its narrative commitment, a film clip exhibits a number of features:
- characters, with actions carried out to achieve some goals,
- objects, involved in such actions or in unintentional events occurring in the clip,
- environments and places, where the action is located,
- links to other film clips in terms of causal connections and story advancements.

Current technology for storing and accessing film clips is mainly based on clip metadata (title, authors, etc.) and users' tags. On the web, repositories such as YouTube, Vimeo, or Anyclip, generally mention the title of the whole film, the list of the involved actors, some specific object manipulated in the scene, the emotional impact on some user (e.g., "best sex scene"), but rarely describe what really happens in the scene and how it is linked to other clips of the same film. Offline, teachers and students surf sequentially through DVD chapters for displaying specific clips or extracting clips for building own repositories; chapters are logical segments thought by DVD editors, that have a loose connection with the original narrative structure. In production, clips that originate from the production process are annotated with the scene code as indicated in the screenplay and the take number, but they do not contain any reference to characters or actions; in the continuation of the production process these data are lost and are (re-)introduced by the film editor's annotation of the clips.

Recommendations also rely on this type of metadata and tags: for example, the "You Tube" engine relies on such data to suggests other clips. Recommendation engines do not take into account the actual content of the clips and such selected excerpts are not related to the narrative content of the film in its entirety. The

retrieval of several clips from the same film rely on the users annotation, e.g. of the type 5/8 for the 5th clip out of 8, so it is not reliable for systematic access.

The project CACHE-22 aims at supporting the access to cinematographic heritage through a semiautomatic semantic annotation of the narrative features of film segments, accompanied by retrieval tools that leverage this annotation. Along with initiatives that aim at categorizing cultural products and providing electronic access to cultural heritage indices, e.g. Europeana (http://www.europeana.eu) or KulturiSampo (http://www.kulttuurisampo.fi), this project addresses the internal structure of items to create advanced search and retrieval tools that grant direct access to the content of the indexed items.

This structure, somehow accounted for in multimedia descriptors, such as MPEG-4 and MPEG-7, has still to be exploited in the tools for accessing the cultural resources. This theme is directly linked with the paradigm of open data, for the free usage of public metadata about cultural and scientific items.

In particular, the project concerns the immaterial audiovisual items of the cinematographic culture. These items are modular in nature and there is a general interest in accessing specific segments of a complex audiovisual or multimedia work. A quick visit to video repositories (such as YouTube) reveals that people upload and search for fragments of large items, for several purposes.

The project works through an integration of automatic analysis of multimedia signals with existing resources for semantic annotation, in order to connect the low-level features of the audio/video signal with high-level structured descriptors for domain independent narrative content. The project realizes a complete technological pipeline, from the analysis of multimedia signals for boundary detection and scene analysis and the integration of ontologies for the identification of meaningful semantic descriptors, to the semi-automatic annotation of narrative descriptors onto video fragments and the implementation of search functionalities based on reasoning over the semantic descriptors. By annotating a number of items (a golden standard corpus), filtered and revised by human experts, the project also envisages the exploitation of the core annotated material through services provided by film heritage institutions and opens to the bootstrapping of a web community for the annotation of large repositories and the creation of learning resources.

The CACHE-22 project has the goal of annotating a narrative media object (typically, a feature film of a cinematographic heritage repository) with a machine-readable language that addresses the internal structure of the object according to its story content. The annotation results from a combination of:
- the reference to light-weight ontology (possibly on a large commonsense ontology), to share the annotation terms and converge on the same semantic descriptors in case of similar media objects;
- an ontology that represents the dramatic features of a media object, that is characters, objects, goals, actions, events, environments;
- automatic approaches over low-level signal features for
  - the detection of clip boundaries that are meaningful from the narrative viewpoint,
  - the detection of characters within clips (through face detection and recognition),
  - the detection of meaningful verbal and written communication within a clip, and
- the connection between the low-level features of media objects and the semantic descriptors;

The annotated film becomes open to advanced retrieval and searching, useful for research purposes and for access by general audience. Specific interfaces can be designed for different types of audiences, in order to encourage the fruition of film inside and outside the exhibitions and collections: for example, finding similarities among stories, comparing different adaptations and remakes, searching similar scenes in a data base of films based on the narrative content.

This approach will achieve the twofold goal of online retrieval of single excerpts and the connection of a clip to the rest of the story; an intelligent search engine can retrieve excerpts on a content base, also inferencing possible similarities through semantic descriptors; the annotation can be the base for recommendation engines that work on the story content, rather than clip metadata or users' tags, and for navigation interfaces that provide an intelligent access to the narrative structure of the content (e.g., taking the viewpoint of a certain character – e.g., "show me the actions carried out by the character Eve along the story of Hitchcock's movie "North by northwest" ").

Activities that can benefit from the result of the CACHE-22 project are:
- the teaching of film history and dramaturgy: narratology, as well as drama studies, can benefit from

a learning resource that can deconstruct the whole media object into meaningful annotated segments;
- story-based film navigation: narrative annotation offers an advanced way to access film and media content to professionals and teachers, who need a more structured way to navigate filmic works than the chapter sequence provided by DVDs;
- search services offered by film repositories or video stocks: more effective search keys than traditional tag or keyword-based search improve the access to cultural resources for the general public, and the professionals' access to film stocks; these services can be generalized to accompanying visual data (film posters, movies, teasers) taken from the open data cloud;
- cultural heritage promotion and dissemination: annotating the story content in film archives supports the design and implementation of interactive exhibitions about the cinematographic heritage;
- cultural offer at large: annotating events, characters and locations in films, enables the creation of heritage exploitation for touristic purposes: for example, exploring places based on their relations to the cinematographic heritage (film-based tours).

Use case scenario 1.
Access to cultural heritage: searching clips in a cultural repository
The visitor of a museum of cinema, in situ or through the web site, wants to access the film repository for retrieving a particular scene of a film. These sorts of memories are very common, as witnessed by the number of scenes extracted from feature films that are present in public repositories, such as You Tube; there are very famous scenes, such as the resolving scene of "Blade Runner" – "I've seen things …" - or the German battle scene in "Gladiator" – "At my signal, unleash hell", often uploaded several times, and less famous scenes, which are meaningful for a few (though still thousands, in many cases!). The museum wants to offers a service for scene retrieval to visitors through some workstations inside the museum as well as to registered visitors through the web site. Moreover, the museum wants to promote the knowledge upon its repository, letting visitors come to know and access less famous items of its collection. The CACHE-22 system recommends scenes that match users' search terms, through an index of film clips based on the narrative ontology; under specified searches can be completed through the narrative templates and the new recommendations will improve knowledge about the cinematographic cultural heritage.

Apart from feature films, video documents held by cultural and government institutions form the memory of the XX century (see, e.g., http://www.criticalpast.com/). Normally these videos are searched and indexed as events "Queen's wedding", "Hitler speech", etc.. Nevertheless a visitor of such repository could be interested in browsing this heritage according to the role of certain type of agent ("athlete"), goal ("winning"), feelings ("happy"). The CACHE-22 annotation can empower this type of content exploration.

Use case scenario 2.
Education: comparing cross-media realizations of a story
A high school teacher wants to describe Shakespeare's "Romeo and Juliet" to her students and wishes to grasp their attention with some examples of how the text has been interpreted in performing arts. After introducing the storyline and the historical coordinates for the original text, she picks up some climactic scenes from the story and asks the students to describe them. Then, she inputs the description into the CACHE-22 system and retrieves a number of videos:
- Romeo+Juliet, 1996 (Baz Luhrmann)
- Romeo and Juliet 1968 (Franco Zeffirelli)
- a recording of the ballet "Roméo et Juliette" (Sergej PROKOFIEV), ballet of Malegot Theatre (Leningrad), 1976

The software interface shows the thumbnails of the selected scene for each of the three videos, each posited on a single timeline that represents the story as told in that video.
The three storylines are aligned according to the matching scenes (each described with a simple sentence, i.e., "Juliet dies", "Juliet drinks the poison", etc.), so the students immediately notice that a scene is missing from the ballet, but also that the final scene is very lengthy in the ballet with respect to the movies. By doing so, they spontaneously flesh out some observations about the different language of cinema and dance.
Then, they watch the three versions of the same scene, and notice the differences in the presentation of the character of Romeo. So they search for the parts played by Romeo in Franco Zeffirelli's film: since too many scenes are selected, they start searching specific scenes by short descriptions of the scene, guided by the teacher (e.g. "the scene in which Romeo kills Tybalt", etc…).

Use case scenario 3
Cultural heritage promotion: reuse of segments in summaries and remixes of various kinds
A cultural institution (e.g., a national archive) wants to promote a retrospective on a director of comedies. Films are annotated according to the CACHE-22 system; so, they are split into clips that are annotated for the participating characters, their actions, the events that occur, the objects manipulated and the environments where action takes place. The retrospective should be better promoted with a video that edits a remix of a number of clips from the films, but the selection of the clips is a hard task. The curator of the retrospective posts some themes of the films into the CACHE-22 search engine to retrieve a number of clips; then, the clips are edited through criteria of action, object or environment similarity. The launching of several retrospectives in a month makes the selection of clips to be exhibited as promo a time consuming task, especially for an institution that have no video editors in their staff, and professional editors could be hired only on particular occasions.

1.2   Progress beyond the state-of-the-art

Describe the state-of-the-art in the area concerned, and the advance that the proposed project would bring about. If applicable, refer to the results of any patent search you might have carried out.

The CACHE-22 project concerns the automatic annotation of narrative media objects oriented to the exploitation of cinematographic cultural heritage. As a research problem, content based access to narrative audiovisuals is situated at the junction of different disciplinary areas, including audiovisual signal analysis, multimedia indexing and retrieval, knowledge representation and reasoning. For this reason, although the current state of the art in each of the involved disciplines has made enough progress to make a joint research effort viable and promising, no significant research efforts can be identified so far in the annotation of and access to narrative content of audiovisuals. So, in the following, we describe the state of the art of the relevant disciplinary areas, with the goal of showing the feasibility of the project based on these results.

We see a number of interdisciplinary issues as crucial to the project goal:
- signal analysis, for the application of algorithms for clip segmentation and scene analysis;
- multimedia indexing, for the annotation of audiovisual material with semantic descriptors, filling the so-called semantic gap between low-level signal features and high-level concept descriptors, is an active research topic;
- ontology engineering, for the design of a suitable ontology for representing narrative features (such as characters, actions and events) thus providing an internal structure to complex narrative (e.g. feature films);
- interface and interaction design, for the exploitation of media indexing for retrieval of sub-items and their presentation in forms that are useful for access to cinematographic cultural heritage.

The project will provide the following advances related to the specific sub-tasks of annotation, search and retrieval for narrative media objects:
- the integration of tools for the automatic analysis of the low-level features of the audiovisual signal (automation of scene detection and boundaries) and methods for the semantic classification of the content;
- the use of large commonsense ontologies for the representation of semantic descriptors in unrestricted domains;
- the representation of the internal structure of narrative media objects and the connection to the encoding of media items.

Video analysis state of the art and beyond

The current state of the art in video analysis and annotation includes several techniques for video decompositions to shots and scenes, as well as techniques for concept detection.

Concerning video decomposition to shots, several techniques exist that achieve good results on a variety of video content have been reported; these results are deemed by the relevant community as being sufficiently accurate for any practical application (Smeaton et al. 2010). Scene segmentation, on the other hand, is still an open research problem, with most approaches of the literature failing to take into account the semantics of the content in performing a task that by definition is based on semantic criteria: different consecutive parts of the video are assigned to the same scene, according to the literature, simply because they present similar low-level audiovisual properties, whereas it is much more than such low-level properties that make humans recognize (and request to consume, in applications such as retrieval) different scenes in a video. This treatment of the scene segmentation problem is particularly problematic when considering cinematographic applications, where it is exactly the semantics of the video that drives the understanding and appreciation of it by its viewers. An attempt to remedy this common deficiency of scene segmentation approaches was recently reported in (Sidiropoulos et al. 2011). Still, though, the latter method continues to disregard sources of information that are particularly important for cinematographic content, such as actor (re-)occurrences.

Concept detection in video is a popular research topic that is currently being investigated extensively. This is grounded on the certain impact that automatic and reliable concept detection would have on a variety of semantics-based multimedia applications. For large-scale concept detection in video, the techniques currently

being used rely mostly on evaluating one or more static, representative keyframes of every video fragment (every shot, in most cases), e.g. see (Smeaton et al. 2009). This is dictated by the computational complexity of treating the video as video, and thus exploiting rich information about the temporal change of the visual signal within every shot. Many techniques of the latter category have been recently developed, e.g. (Mezaris et al. 2009)(Niebles et al. 2008), but their computational complexity often renders prohibitive their practical use in real-life applications, or requires that several simplifications to the actual video description method are introduced.

The problem of event detection, particularly considering complex high-level events rather than elementary actions, is not so well-studied; research on high-level event detection from video is still in its infancy.

State of the art in automatic analysis of the low-level features of the audiovisual signal for clip segmentation (automation of scene detection and boundaries) and semantic classification with building of large commonsense ontologies for the representation of semantic descriptors

The coverage of the semantic gap in video indexing and retrieval has gone through a continuous increase of the vocabulary of high-level features or semantic descriptors. Starting from a few tens of the first TRECVid conferences, semantic descriptors now amount to a few thousands (http://www.lscom.org). As concepts increase in number, the search task benefits from the creation of semantic relations over individual concepts. In fact, related concepts can reinforce each other through the mutual influence of their detectors, and even the use of low accuracy detectors can improve search results substantially, provided that the lexicon consists of a large set of concepts.

The incorporation of semantic relations has led to the creation of ontologies, to organize hundreds or thousands of concepts. LSCOM is a concept ontology targetedly designed for a corpus of broadcast news (LSCOM 2006); the MediaMill dataset relies on a set of 101 semantic descriptors that are best suited for that repository (Mediamill 2006). In this sense, standard large scale resources contribute to make descriptors as interoperable as possible, so the results and relations devised for one application can be shared with other applications. Also, there are approaches for the acquisition of rules that define complex semantic concepts, starting from initial, simple rules and exploiting existing semantic relations (provided by external resources, such as light and foundational ontologies) (Ballan et al. 2010).

The effort in the creation of semantic descriptors has the goal of best meeting the queries of users, in order to be effective in the search task, see the TRECVid story in (Smeaton et al. 2009). In this task, the manual or semiautomatic generation of annotations is crucial in providing ground truth data for evaluation purposes and as training data to support knowledge acquisition and learning (Nixon et al. 2011).

State of the art in the representation of the internal structure of narrative media objects and the connection to the encoding of media items

Most work in this field is oriented to content generation, customization and reuse, assuming a production perspective to the representation of media objects. According to this perspective, a narrative object consists of a number of segments (or sub-objects), collected in a repository, from where they are retrieved for editing purposes. Different internal structures have been used as the base for different annotations. For example, the Strata system (Smith, Davenport 1992) allows the user to associate concepts to, potentially overlapping, spans of video. More recent work has combined the high-level view of audiovisual structures with the lower level view provided by standards such as MPEG, resulting in frameworks such as the Core Ontology for Multimedia, COMM (http://comm.semanticweb.org/, (Arndt et al. 2007)) that allow the association of concepts from an existing vocabulary with specific areas and time-fragments of any media type. Similar work is surveyed by (Dasiopoulou et al. 2010)).

Together with annotation methods and editing programs, there are a number of complete systems developed with artistic goals or specific genres in mind. The application Soft Cinema (http://softcinema.net, 2005–2011) hybridizes the paradigms of cinema and human-computer interaction, and enhances the new media practices in video production. In Soft Cinema, media elements are stored in a large database, and the narrative is generated through the selection and editing of clips (a 'database narrative' according to

Manovich (Manovich 2001). Each clip is assigned keywords that describe both its "content" (geographical location, presence of people in the scene, ...), and its "formal" properties (i.e., dominant color, contrast, camera motion).

In the field of interactive TV, ShapeShifting TV (Ursu et al. 2008) is a system for generating interactive TV narrative that encompasses both the annotation of the media items, and the formal description of the narrative structure, the latter accomplished through the Narrative Structure Language. This system was employed for the production of interactive tv movies ("Accidental Lovers", 2006) and documentaries ("A Golden Age", 2007).

The Narrative AbstractModel (Jung et al. 2007) is targeted at the summarization of narrative audiovisuals. This model splits the video into segments, with algorithms based on stylistic elements (such as shot and scene detectors), informed by a manually built story structure. Then, based on this representation, it computes video abstracts for TV review services of soap operas. During this process, the soap-opera episode is represented as a graph of interconnected narrative nodes, with an evaluation of the degree to which each interconnection realizes a story progress; the intelligent component of the program detects the sequence of narrative units that mostly contribute to the progress of the story (Degree of [story] Progression—DoP—measure), generates a graph that encodes all the possible sequences of connected narrative elements, and selects the sequence (path) with the maximum DoP to propose an abstract.

Finally, there two recent projects address the annotation of video and its structure. The Advène project (Richard et al. 2008) addresses the annotation of digital video, and is not specifically targeted to the narrative content. In Advène, the fragments of the video are annotated with free textual description of the content, cross-segment links, transcribed speech, etc. This information can be exploited to provide advanced visualization and navigation tools for the video. For example, as a result of the annotation, the video becomes available in hypertext format. The annotation is independent from the video data and is contained in a separate package that can be exchanged on the net.

Drammar is an ontology-based model for an annotation schema of the narrative features of media objects (Lombardo, Damiano 2011). Though retaining the core assumptions of most narratological theories, Drammar provides a theory a-specific description of the narrative content of media objects. Such a model is aimed at building large corpora of narrative media (with no specific constraint on format), that can be employed to build and test specific models and applications in the areas of access to cultural heritage and learning. Drammar will be the base for the CACHE-22 project.

State of the art in media ontologies, with links to narrative ontologies

Early work on composing narratives from video fragments was done in Media Streams (Davis 1995), that included an engine that allowed users to construct narrative video sequences based on querying the rich annotations of the video database. A significant contribution to the design of a narrative ontology comes from the field of interactive digital storytelling, i.e., the on-the-fly generation of narrative audiovisuals by taking into account the input from the users. There are two broad categories of interactive storytelling systems: story–based appraches (e.g., (Riedl, Young 2006)) and character–based approaches (e.g., Pizzi et al. 2007).

Story–based systems are characterized by centralized architectures, in which the system is driven by the unifying principle of a story. Story–based architectures tend to incorporate sophisticated story models to account for the structural aspects of narration (e.g., Szilas 2003). Character–based systems rely on the autonomous behavior of characters, letting characters interact in a constrained environment so as the emergent situations can be interpreted as narrative structures. This approach has been encouraged by the availability of conceptual and practical tools that implement the characters' deliberation through the notion of intelligent agent (Theun et al. 2003).

The Façade system (Mateas, Stern 2005), an interactive drama performance, adopts a mixed strategy by encoding multi-agent plans for the characters (included the user, seen as a character) while at the same time controlling the story advancement through numerical parameters that represent the so-called "story value".

The Dramatour system (Lombardo et al. 2008) uses metadata to annotate character-based narrative unit that are edited on the fly for visitor guidance in a museum setting.

The Vox Populi system (Bocconi et al. 2008) uses the Toulmin model of argumentation to provide a structure for automatically combining multiple fragments into coherent treatises on a topic selected by the user. The system combined 3 sets of descriptors for the domain knowledge (a small, structured domain-

specific vocabulary selected for the corpus available), for the role in the argument (using elements identified in the Toulmin model) and for filmic annotations (such as camera movement and framing) that ensured that the combined fragments were visually coherent.

State of the art in technological projects for the exploitation of cinematographic cultural heritage in both access and learning

There are a few projects on the exploitation of the cinematographic (and audiovisual) cultural heritage. These projects explore the use of innovative access techniques (e.g., visual interfaces) for access to cultural heritage, with goals of promotion and dissemination. Some of them, like ACMI are specifically targeted at learning, while some others specifically address dissemination and fruition (AVIE). Finally, CINeSPACE and Linked Movie Database take film heritage and archives as their domain. In the following, we list the most relevant ones.

SP-ARK (http://www.sp-ark.org)
SP-ARK is an interactive online project based on the multi-media archive of film-maker Sally Potter.
A unique educational resource, SP-ARK is designed to enhance your access to, and knowledge of, film and filmmaking, whatever your interests.
By telling the story of a film by showing every aspect of the production –from the initial idea through screenwriting, budgeting and casting to the composition of an individual shot – SP-ARK invites you to become part of a new generation of film scholars and viewers who understand film inside out.

ACMI Generator (http://generator.acmi.net.au)
http://www.museumsandtheweb.com/mw2011/best/education/acmi_generator
ACMI Generator is an online creative studio (website and interactive officially launched in September 2010), that seeks to educate and encourage a vibrant young audience to develop strong storytelling skills in all screen mediums. The major motivation is that, while the creation of media is ubiquitous amongst young people, the principles of storytelling are often lost in the rush to upload the latest video and share with peers. The design of ACMI Generator is intended to encourage reflection on what makes a good story – a story that really works – and to provide a platform for students to gain expert knowledge and collaborate and share that knowledge with others. Generator is a highly engaging social media platform, where students and teachers can share creative work and education resources, downloading from, and uploading to, the extensive and free library of film and other media.
The Generator Gallery comprises more than 100 diverse stories which can be rated, shared via email and social media, selected as favourites, commented on and tagged and downloaded. The Free Media Library (governed by a Creative Commons Attribution) consists of over 1000 individual copyright free items (video, images and audio files) that are able to be downloaded, re-used, adapted, remixed, mashed up and republished for educational purposes. Students can upload their own media to the library to share with the Generator community.
Most importantly Generator is innovative, providing a connection between current technologies and classroom practices, but also inspiring new ways to think about, to engage with and to share creative content and practices. ACMI Generator takes learning beyond the classroom and connects students with peers, professional practitioners and diverse creative communities.

AVIE system
Based on the consideration that "creative production overtake basic access as the primary motive for interaction with databases, archives and search engines (Manovich 2008)", the AVIE system enables immersive visualization modalities (full body immersion, stereoscopy, spatialized sound and camera-based tracking) as a way to support reuse and participative culture in the field of digital humanities [Kenderline and Hart 2011]. This practice, termed "cultural data sculpting", constitutes an important attempt in the use of new media technologies to make large and heterogeneous collection meet the needs of specific institutions and audience types, by using mainly visual tools
(http://www.museumsandtheweb.com/mw2011/programs/cultural_data_sculpting_omni_spatial_interac).

Cinemetrics

An attempt to develop tools for the visualization and fruition of signal data in movies is given by the Cinemetrics project (http://cinemetrics.fredericbrodbeck.de/) is an independent, open source project for measuring and visualizing movie data, by creating a visual "fingerprint" of the movie. The analysis can be extended to the media items related to movies, such as posters or pictures.

Linked Movie Database
In the filed of semantic annotation, the Linked Movie Database (http://www.linkedmdb.org/) aims at publishing the first open semantic web database for movies, including a large number of interlinks to several datasets on the open data cloud. Mainly aimed at interlinking the movie domain with other semantic resources, such as DBpedia, WordNet or MusicBrainz, the Linked Movie Database relies on the use of a tool, ODDLinker, to finding links between different data sources.
While the project does not encompass specific visualization tools, it is compatible with semantic browser, and constitutes a semantic middleware onto which large scale applications for smart film archives can be created.

SHOTS: SHared annOTationS on multimedia content (http://www.filmforumfestival.it/?p=1120)
The project SHOTS aims at creating a set of methodologies and tools for feature extraction, analysis, annotation, and retrieval of audiovisual content. The final goal is to provide a framework to carry out on-line film analysis, to interact with multimedia content, and to share the analysis results with a community of users.
The main application of SHOTS is as a tool for film scholars and, in particular, those involved in film analysis aimed at scientific research and teaching. Moreover, it can be exploited also by media experts and amateurs, who can improve their interaction with audiovisual content.
A further aim is the development of manual annotation tools, to supplement and correct the automatic annotations. With the same tool it will be possible for final users to track the results of scientific research directly on audiovisual content.
SHOTS platform will help final users to interact more deeply with the art of cinema and, at the same time, allow the sharing of filmic analysis with user-generated content, but also with the studies of experts and scholars of the scientific community.

CINeSPACE, http://cordis.europa.eu/ist/digicult/cinespace.htm
Ontologies are used also in CINeSPACE, an European research project (ended 2009) that aims at designing and implementing a mobile rich media collaborative information exchange platform, scalable, accessible through a wide variety of networks, and therefore, interoperable and Location-Based for the promotion of Film Heritage, going beyond the current state of the art. CINeSPACE enables users to interact with Location-Based multimedia contents while navigating a city. Audiovisual information are be delivered through a unique and portable low-cost wireless high definition near-to-the-eye display and audio phones, and also comprise a small camera able to record or send what the user is "seeing". This information can be uploaded to a database through a WLAN hot spot or a 3G connection in order to create collaborative experiences with other end users. Hence the project is using films as part of Cultural Heritage and extracting, delivering, and displaying for the user, on site, and, in order to do so, it integrates existing standards (ontologies, MPEG) to provide semantic access to indexed multimedia cultural content.

European projects related to CACHE-22

There are a number of European initiatives that are related to CACHE-22 : on one hand, we have technological projects (as is the case for CACHE-22), that make CACHE-22 viable because of the maturity of the technologies involved; on the other, we have cultural projects, that provide inspiration for CACHE-22 scenarios.

LinkedTV project (http://www.linkedtv.eu/)
In an era where television content and Web content are seamlessly connected, and the distinction between broadcaster becomes irrelevant for the final user, Television Linked To The Web (LinkedTV) provides a novel practical approach to Future Networked Media. It is based on four phases: annotation, interlinking, search, and usage (including personalization, filtering, etc.).
CACHE-22 is different from LinkedTV in its application to narrative media and the internal structure of

media with respect to its content; LinkedTV is geared towards the technological achievements of a common cross-platform between TV and Web, while CACHE-22 focuses on content and its interpretation for usage and repurpose. CACHE-22 can benefit from the results of the Linked TV project in the exploitation of the methods devised for annotation, interlinking, search and usage. Methods will be extended to the internal structure of media and specified for the narrative character of media, including the detection of macro-features, such as genre (e.g., Comedy, War, …) and the linking of characters' traits across media items (e.g., James Bond's features through 007 movie series).

ANSWER project (http://www.answer-project.org)
The goal of the ANSWER project is a notation system for describing the creation of multimedia content, a bridge between digital media production and animation for game design e previsualization. The ANSWER project has developed a "Director Notation", a language for the symbolic description of the author's creative intent, and tools for the automatic generation of animated pre-visualisations for film and game pre-production from this notation. An interface allows the Director to work imaginatively with the notation as it is authored, and an iterative scheme is introduced to ensure that developments of ideas and the realities of the actual production are reflected in the final notation description. The major innovations brought by the ANSWER project are:
- a film production ontology, i.e. a semantic model of the film production process;
- a formal notation system for the representation of the direction issues in film and game;
- an application for the automatic generation of animated 3D 'pre-vis' storyboards of films and contents of games from the machine-processable representation;
- an application for the automatic extraction of metadata with goals of content-based retrieval of raw footage during post production;
- synchronization of the user's notation input to what is actually filmed;

Differently from ANSWER project, CACHE-22 is not geared to the audiovisual production and the annotation of film clips concerns the narrative qualities rather than the expression of the director's intentions towards the staging of characters and scene layout. Instead, CACHE-22 is intended to abstract from actual mise en scene and staging details, as its main strength in the access to narrative content. CACHE-22 can benefit from the results of the ANSWER project in the annotation of the spatial (layout) aspects of a scene and the physical qualities of the action occurring therein.

IM3I - Immersive Multimedia Interfaces (http://im3i-server.hku.nl/)
This project has developed intelligent and flexible repository software, for supporting the work of content holders in producing, retrieving and repurposing digital artifacts. Im3i platform analyses media artifacts on-the fly on both physical and semantic characteristics and provides a comprehensive authoring environment which allows non-technical users to make the most from their media collections, whether as part of in-company production workflow or in repurposing collections to distinct end user groups.
Based on rule learning to create new structured concepts from existing ones, this project shares with CACHE-22 the use of ontologies to describe and access to digital artifacts. Differently from CACHE-22, it not geared on audiovisuals only and is not aimed at modeling the narrative features of multimedia content. However, the CACHE-22 project will benefit from the results of its mixed initiative approach to annotation, since the project has devised specific tools for supporting the role of the user (and the community of users) in the annotation and search.

NM2 / New media for new millenium (finished in September 2007) The Future of Media Production
NM2 was a collaborative research project uniting leading creative and technology experts from across Europe to address a great opportunity for businesses and consumers: how to develop compelling new media forms which take advantage of the unique characteristics of broadband networks. NM2 was about creating a variety of new media genres using all of the facilities of modern 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. The design of an explicit story model and its use in authoring tools is a main asset output by this project, that will affect the design of the narrative model in CACHE-22.

TA2 (Together Anytime, Together Anywhere) 2008-2011
"Making communications and engagement easier among groups of people separated in space and time."
The project TA2 fosters family relations by developing innovative applications that rely on new media. Since communication devices and new media mostly address the needs of individuals, the project aims at providing services that allow groups of people, linked to each others by family relationships, to experience group-oriented communications and experiences: "Through the TA2 system, stories are automatically generated from home-related content, the personal home video or from the antics of a lively game".
The project encompasses the design of five media experiences, ranging from automatic capture and editing of audio-visual content to the storytelling (described as reading a book together while being apart). Innovations consist of a language to describe interaction scenarios, multimedia composition and transmission technologies, automatic framing and segmentation tools given a scenario, simplified editing of media based on narrative templates, and the design of new experiences based on new media (like gameplay, social media, etc.).
TA2 shares with CACHE-22 the objective of coupling low level signal analysis (face recognition, segmentation, etc.), with some higher level organization of the content, such as communicative and narrative structures, but is mainly oriented to the exploitation of media content for social and communicative purposed. It acknowledges the role of semantic descriptors in audiovisual content, but it does not ensompass the semantic representation of the narrative structures. Rather, it aims at selecting and adapting multimedia resources to social interaction scenarios characterized by a virtual shared space. So, TA2 privileges editing languages for multimedia such as SMIL, through which recognized concepts are shaped into customized, communicative content. Differently from CACHE-22, it does not focus on search and retrieval but on intelligent and semi-automatic editing, presentation and delivery of media resources in social communicative scenarios.

AXES (2011-2014):
Goal of AXES is to develop tools that provide various types of user with new engaging ways to interact with audiovisual libraries, helping them discover, browse, navigate, search and enrich archives. NISV is one of the three content providers. In particular, apart from a search-oriented scheme, the project will explore how suggestions for audiovisual content exploration can be generated via a myriad of information trails crossing the archive. This will be approached from three perspectives (or axes): users, content, and technology. Along these axes, open source solutions will be tested onto use cases, validated and integrated in a platform of services.
While AXES is oriented to services for multimedia analysis, CACHE-22 assumes the integration of these services into a narrative oriented annotation pipeline, geared to filmic heritage. For this reason, AXES may provide useful tools, within its platform, for the integration is CACHE-22, and a confirmation of the experimental validity of these tools in the pervasive domain of narrative.

ADVANCEMENTS WITH RESPECT TO STATE OF THE ART

The CACHE-22 project has the goal of bridging the semantic gap in the case of narrative media objects by providing an ontology for the representation of the narrative features and the attachment of the content representation about the actions/events occurring and the characters/objects/environments involved to a narrative template.

Integration of low-level and high level features in narrative audiovisual analysis

Narrative audiovisuals are unrestricted in domain and require a great variety of semantic descriptors, with complex descriptors realized through concept relations, thus constituting a valid test for the application of semantic technologies. Moreover, the structure of narratives offers an appropriate testbed for the annotation of several layers of descriptors, depending the interpretation of and for establishing a causal connection over several sub-objects.
Consider, for example, a scene from the classic Hollywood film "North by Northwest", by Alfred Hitchcock, taken from YouTube (http://www.youtube.com/watch?v=8sKTO32r5Zs). In this scene, set on the train which takes the main character, Roger Thornill, from New York to Chicago, an undercover spy, Eve

Kendall, seduces Roger to manipulate him. On the YouTube video we consider, this scene is named after the main event occurring in it, namely the kiss between the Roger and Eve - North by Northwest Kiss Scene. The scene is shortly described as "Cary Grant and Eva Marie Saint kiss on the train"; interestingly enough, the tags attached to it ("Cary Grant Eva Marie Saint North By Northwest Hitchcock classic film") do not refer to the narrative content, but only encompass production data, such as actors, directors and title, accompanied by a value statement by the user ("classic film"). The "kissing scene" can be described as the action of "kissing" executed by the two characters Roger and Eve; the "kissing" action itself also encompasses the chatting action that intersperses the kissing, so some annotator would describe the same scene as "flirting", working at a higher level. This finally, if we extend the scene boundaries, the appropriate descriptor would be "Eve helping Roger in escaping from police", revealing the goal of the character Eve achieved with the action of hiding Roger, and at a yet higher level, the audience realize Eve is actually seducing Roger, obeying to the commands of a gangster.

So, if we consider the decisions taken by a human annotator in providing a semantic classification of a scene, we can observe two facts.
- The detection of scene boundaries depends on the scene interpretation, though it can be supported, in some cases, by referring the original screenplay (actually, the analysis of the differences between the screenplay and the actual film is an interesting learning case in cinema studies); conversely, the annotation depends on the human or automatic interpretation of the intentional actions and the unintentional events that form the observable incidents and the capability to identify the semantic concepts (through natural language terms for humans filling the semantic gap for machines).
- There exists a causal/spatial/temporal connection over the incidents occurring in a narration; each unit or film clip has a role in the narration and the whole story features an infrastructure for the accommodation of units in the global context of the narration. Hierarchical relations as well as long-distance relations can be established over the units, linking the characters' desires with the actions they carry on and the events that occur unintentionally in the story world.

The design of an integrated annotation pipeline will allow the CACHE-22 project to integrate the audiovisual features extracted from the low level signal analysis, and the concept automatically detected from these features with the high level semantic model of the narrative content. Through the annotation tools, human annotators will be able to explore the mapping of low level features onto higher level semantics, modifying this mapping when needed and accommodating the high level features into the narrative model. Shot and scenes will become narrative units, recognized entities will become characters, object and locations, labeled according to large scale resources, in a intentional event chain that bridges the gap between story and signal. Building on the ontology–based model called Drammar, CACHE-22 will produce a structured information with respect to the event structure and to the storyline, that can represent both individual media items and items connected through a relations in the complete story line.

Audiovisual signal analysis.

We will extend and adapt state of the art shot and scene segmentation to cinematographic content, taking into account information that is central to the nature of this content, such as actor occurrences (derived by means of automatic face/person detection and unsupervised face clustering). In addition to this, we will also detect and exploit finer-granularity temporal segments of the video, namely sub-shots, which can also be useful for video content re-use.
Concerning concept detection, we will build on the current state of the art in order to develop techniques that will take into account some form of temporal information, so as to feature improved accuracy, and at the same time exhibit computational complexity that will be comparable to that of keyframe-based approaches. At the same time, we will consider the use of computationally efficient learning methods in combination with these temporal features. Our concept detection advances will also help us in developing effective techniques for higher semantic level event detection from video, which will rely on concepts. Such approaches are shown in very recent activities, such as the TRECVID Multimedia Event Detection (MED) task to be best-suited for going beyond simple action/activity recognition in video (e.g., for detecting if one or more actors/objects are engaged in complex interactions, instead of detecting simply if a person is running or not).

Annotation of narrative media

The paradigm of linked data (Heath, Bizer 2011) offers a way to link a media object with external semantic resources. In the World Wide Web, classes, properties and individual of any ontology can be referred anywhere by using URIs to identify them. Thanks to this mechanism, the semantic annotation of a media object can rely on an external ontology as a terminological base without requiring an explicit integration of it in the annotation model.

In the linked data scenario, media objects are not simply annotated with semantic descriptors, extracted from light scale, domain ontologies, but a meta-model is imposed on them, in the form of the narrative annotation schema, leaving to external, shared ontologies the task of providing the vocabulary for describing the story content. While the meta-model encompasses character-enacted events, the process types these events consist of, and the type the characters belong to, are given by the external ontologies, according to the paradigm of linked data. For example, we can refer to individuals that belong to different datasets through the *type* property. In the triple (\<object,type,URI\>), the value refers to the URI of a concept in another ontology, e.g. YAGOSUMO (a commonsense knowledge base (De Melo et al. 2008), merging the Suggested Upper Merged Ontology - SUMO (Pease 2002) and the Yet Another Great Ontology - YAGO (Suchanek 2007)). So, if the object is a car, the type property of this object will take as its value the URI of the concept of "car" in YAGOSUMO.

YAGOSUMO incorporates almost 80 millions of entities from YAGO (which is based on Wikipedia and WordNet) into SUMO, a highly axiomatized formal upper ontology. This merged ontology provides very detailed information about millions of situations, including entities, processes, and events that can be positively used not only for annotation purposes, but also for automated knowledge processing and reasoning.

Bibliography

Arndt R, Troncy R, Staab S, Hardman L, Vacura M (2007) COMM: designing a well-founded 1076 multimedia ontology for the web. The semantic web, pp 30–43.

C. Baker, C. Fillmore, and J. Lowe. The berkeley framenet project. In Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics-Volume 1, pages 86-90. Association for Computational Linguistics, 1998.

L. Ballan, M. Bertini, A. D. Bimbo, and G. Serra. Video annotation and retrieval using ontologies and rule learning. IEEE MultiMedia, pages 80-88, October-December 2010.

L. Bentivogli, E. Pianta, and C. Girardi. Multiwordnet: developing an aligned multilingual database. In First International Conference on Global WordNet, Mysore, India, 2002.

M. Bertini, R. Cucchiara, A. Del Bimbo, and C. Torniai. Video annotation with pictorially enriched ontologies. In Proc. of IEEE International Conference on Multimedia & Expo (ICME), Amsterdam, NL, July 2005.

M. Bertini, A. Del Bimbo, and G. Serra. Learning ontology rules for semantic video annotation. In Proceedings of the 2nd ACM workshop on Multimedia semantics, MS '08, pages 1-8, New York, NY, USA, 2008. ACM.

S. Bocconi, F. Nack, L. Hardman. Automatic generation of matter-of-opinion video documentaries. Journal of Web Semantics 6(2): 139-150 (2008)

D. Brickley and R. V. Guha. Resource Description Framework (RDF) Schema Specification. W3C Recommendation, World Wide Web Consortium, 2000.

Butler S, Parkes A (1997) Film sequence generation strategies for automatic intelligent video editing. Appl Artif Intell 11:367–388.

B. Coyne, O. Rambow, J. Hirschberg, and R. Sproat. Frame semantics in text-to-scene generation. In R. S. et al., editor, KES 2010 Part IV, LNAI 6279, pages 375-384, 2010.

G. De Melo, F. Suchanek, and A. Pease. Integrating yago into the suggested upper merged ontology. In Tools with Artificial Intelligence, 2008. ICTAI'08. 20th IEEE International Conference on, volume 1, pages 190-193. IEEE, 2008.

Dasiopoulou S, Tzouvaras V, Kompatsiaris I, Strintzis M (2010) Enquiring MPEG-7 based multimedia ontologies. Multimed Tools Appl 46(2):331–370).

M. Davis. Media Streams: Representing Video for Retrieval and Repurposing.
PhD thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts,
1995.

A. Ekin and A. M. Tekalp. Automatic soccer video analysis and summarization. In Storage and Retrieval for Media Databases, pages 339-350, 2003.

A. R. Francois, R. Nevatia, J. Hobbs, and R. C. Bolles. Verl: An ontology framework for representing and annotating video events. IEEE MultiMedia, 5:76-86, 2005.

A. Hauptmann. How many high-level concepts will fill the semantic gap in video retrieval? In Proceedings of the ACM International Conference on Image and Video Retrieval, 2007.

T. Heath and C. Bizer. Linked data: Evolving the web into a global data space. Synthesis Lectures on the Semantic Web: Theory and Technology, pages 1-136, 2011.

L. Hollink, T. A. Schreiber, B. J. Wielinga, and M. Worring. Classification of user image descriptions. International Journal of Human-Computer Studies, 61(5), 2004.

A. Jaimes and J. R. Smith. Semi-automatic, data-driven construction of multimedia ontologies. In Proceedings of the 2003 International Conference on Multimedia and Expo - Volume 2, ICME '03, pages 781-784, Washington, DC, USA, 2003. IEEE Computer Society.

Jung B, Song J, Lee Y (2007) A narrative-based abstraction framework for story-oriented video. ACM Transactions on Multimedia Computing, Communications and Applications 3(2):1–28.

Sarah Kenderdine and Tim Hart (2011). Cultural Data Sculpting: Omni-spatial Visualization for Large Scale Heterogeneous Datasets. Proceedings of Museums and the Web, Philadelphia, 2011.

R. Leonardi and P. Migliorati. Semantic indexing of multimedia documents. IEEE MultiMedia, 9:44-51, April 2002.

Lombardo V, Nunnari F, Damiano R, Pizzo A, Gena C (2008) The canonical processes of a dramatized approach to information presentation. Multimedia Systems Journal (Special Issue on the Canonical Processes of Multimedia Production) 14(6):385–393.

V. Lombardo and R. Damiano. Semantic annotation of narrative media objects. Multimedia Tools and Applications, pages 1-33, 2011.

Lev Manovich (2001) The Language of New Media, MIT Press, 2001.

Mateas M, Stern A (2005) Structuring content in the façade interactive drama architecture. In: Proceedings of artificial intelligence and interactive digital entertainment.

V. Mezaris, A. Dimou, I. Kompatsiaris, "On the use of feature tracks for dynamic concept detection in

video", Proc. IEEE International Conference on Image Processing (ICIP 2010), Hong Kong, China, September 2010.

G. Miller. Wordnet: a lexical database for english. Communications of the ACM, 38(11):39-41, 1995.

Nack F, Lindsay AT (1999) Everything you wanted to know about mpeg-7. IEEE MultiMedia.

Nack F, Putz W (2001) Designing annotation before it's needed. In: Proceedings of the 9th ACM international conference on multimedia, pp 251–260.

M. Naphade, J. R. Smith, J. Tesic, S.-F. Chang, W. Hsu, L. Kennedy, A. Hauptmann, and J. Curtis. Large-scale concept ontology for multimedia. IEEE MultiMedia, 13:86-91, July 2006.

J. C. Niebles, H. Wang, and L. Fei-Fei, "Unsupervised Learning of Human Action Categories Using Spatial-Temporal Words," Int. J. of Computer Vision, vol. 79, no. 3, pp. 299–318, Sept. 2008

L. Nixon, S. Dasiopoulou, J.-P. Evain, E. Hyvonen, I. Kompatsiaris, and R. Troncy. Multimedia, broadcasting, and eculture. In J. Domingue, D. Fensel, and J. A. Hendler, editors, Handbook of Semantic Web Technologies, pages 911-975. Springer-Verlag, Berlin Heidelberg, 2011.

A. Pease, I. Niles, and J. Li. The suggested upper merged ontology: A large ontology for the semantic web and its applications. In In Working Notes of the AAAI-2002 Workshop on Ontologies and the Semantic Web, 2002.

Pizzi D, Charles F, Lugrin J, Cavazza M (2007) Interactive storytelling with literary feelings. In: The second international conference on affective Computing and Intelligent Interaction (ACII2007). Springer, Lisbon, Portugal.

G. Prince. A Dictionary of Narratology. University of Nebraska Press, 2003.

Richard, B., Prié, Y., Calabretto, S.: Towards a unified model for audiovisual active reading. In: Tenth IEEE International Symposium on Multimedia. pp. 673-678 (Dec 2008).

Riedl M, Young M (2006) From linear story generation to branching story graphs. IEEE Journal of Computer Graphics and Applications, pp 23–31.

C. Saatho, S. Schenk, and A. Scherp. Kat: The k-space annotation tool. In SAMT 2008, Demo Session Proceedings, 2008.

P. Salembier. Overview of the MPEG-7 standard and of future challenges for visual information analysis. EURASIP Journal on Applied Signal Processing, 2002(4):343-353, 2002.

P. Sidiropoulos, V. Mezaris, I. Kompatsiaris, H. Meinedo, M. Bugalho, I. Trancoso, "Temporal video segmentation to scenes using high-level audiovisual features", IEEE Transactions on Circuits and Systems for Video Technology, vol. 21, no. 8, pp. 1163-1177, August 2011.

A. F. Smeaton, P. Over, and A. R. Doherty, "Video shot boundary detection: Seven years of trecvid activity," Computer Vision and Image Understanding, vol. 114, no. 4, pp. 411–418, April 2010.

A. F. Smeaton, P. Over, and W. Kraaij. High-level feature detection from video in trecvid: a 5-year retrospective of achievements. In A. Divakaran, editor, Multimedia Content Analysis, Signals and Communication Technology, pages 151-174. Springer Science+Business Media, LLC, 2009.

Smith TGA, Davenport G (1992) The stratification system. a design environment for random access video. In: ACM workshop on networking and operating system support for digital audio and video.

C. G. Snoek, M. Worring, J. C. van Gemert, J.-M. Geusebroek, and A. W. Smeulders. The challenge

problem for automated detection of 101 semantic concepts in multimedia. In Proceedings of ACM Multimedia,, pages 421-430, Santa Barbara, USA, October 2006.

F. Suchanek, G. Kasneci, and G. Weikum. Yago: a core of semantic knowledge. In Proceedings of the 16th international conference on World Wide Web, pages 697-706. ACM, 2007.

Szilas N (2003) Idtension: a narrative engine for interactive drama. In: Proc. 1st international conference on Technologies for Interactive Digital Storytelling and Entertainment (TIDSE 2003). Darmstadt, Germany.

M. Theun, S. Faas, D.K.J. Heylen, and A. Nijholt. (2003) *The virtual storyteller: Story creation by intelligent agents.* In: TIDSE 2003: Technologies for Interactive Digital Storytelling and Entertainment, 24-26 March 2003, Darmstadt, Germany. pp. 204-215. Fraunhofer IRB Verlag. ISBN3-8167-6276-X

S. Tonelli and D. Pighin. New features for framenet -wordnet mapping. In CoNLL'09: Thirteenth Conference on Computational Natural Language Learning, Boulder, CO, USA, 06/2009 2009. ACL.

Ursu MF, Kegel I, Williams D, Thomas M, Mayer H, Zsombori V, Tuomola ML, Larsson H, Wyver J (2008) Shapeshifting tv: interactive screen media narratives. Multimedia Syst 14(2):115–132.

G. Zhu, M. Yang, K. Yu, W. Xu, and Y. Gong. Detecting video events based on action recognition in complex scenes using spatio-temporal descriptor. In Proceedings of the ACM Multimedia Conference, pages 165-174, 2009.

1.3 S/T methodology and associated work plan

A detailed work plan should be presented, broken down into work packages (WPs) which should follow the logical phases of the implementation of the project, and include consortium management and assessment of progress and results. (Please note that your overall approach to management will be described later, in section 2).

Please present your plans as follows:
i)   Describe the overall strategy of the work plan *(Maximum length – one page)*
ii) Show the timing of the different WPs and their components (Gantt chart or similar).
iii) Provide a detailed work description broken down into work packages:
   • Work package list (please use table 1.3a);
   • Deliverables list (please use table 1.3b);
   • List of milestones (please use table 1.3c)
   • Description of each work package (please use table 1.3d)
   • Summary effort table (1.3e)
iv) Provide a graphical presentation of the components showing their interdependencies (Pert diagram or similar)
v)   Describe any significant risks, and associated contingency plans

Note: The number of work packages used must be appropriate to the complexity of the work and the overall value of the proposed project. The planning should be sufficiently detailed to justify the proposed effort and allow progress monitoring by the Commission.

i)   Describe the overall strategy of the work plan *(Maximum length – one page)*

The CACHE-22 project aim at defining a technological pipeline, based on the semi-automatic analysis of contents, for the exploitation of cinematographic cultural heritage. In particular, CACHE-22 focuses on drama items, producing an internal structure and identifying significant clips to be annotated and searched on a content basis. Automation of annotation and search is relevant for the massive application of the technology to the heritage archives, and the internal structure as well as the annotation marks are driven by ontologies, that formalize commonsense knowledge, dramatic features, and media objects. The project will be carried out through a thorough integration of academic, industrial and content partners. The management (Workpackage 1) will make possible such a strong interrelation of competences and objectives; this will be reflected by dissemination too, addressing different audiences and exploiting results in various directions.

The whole project starts with the tuning of two ontologies:
1. a media ontology, that incorporates the features of media objects in general, including the Dublin Core metadata for cultural resources;
2. a drama ontology, that introduces an internal structure for the organization of scenes of a dramatic media object and the structure of each scene.

This workpackage (WP 2) is based on previous experiences of two of the partners, CWI and UNITO, respectively.

At the same time, the project addresses the definition of the corpora, that will work as a test for the technology pipeline, and the setting of use case scenarios, that will inspire the design of exploitation applications for access and learning of cultural heritage. This workpackage (WP 3) is mostly carried out by the two content partners, namely NISV and LUCE, with the contributions of the academic researchers and the industrial partners. From the corpora, we will also extract a sample for testing and debugging the beta versions of the applications to be implemented in the subsequent workpackages.

With these two workpackages at good stages of advancement, we start working at the design of annotation guidelines and a suitable interface for realizing a sample annotated corpus of dramatic media objects. We also elaborate a light-scale commonsense ontology, driven by the actual content selected for the exploitation and in collaboration with the partners working on the signal analysis. The implementation of the guidelines through the designed interface will allow human experts to annotate the sample corpus. Content institutions will work under the guidance of research institutions that will provide training and assistance in the

annotation phase. This workpackage (WP 4) will produce the sample annotated corpus that will be the golden standard for inspiration and evaluation of automatic phases.
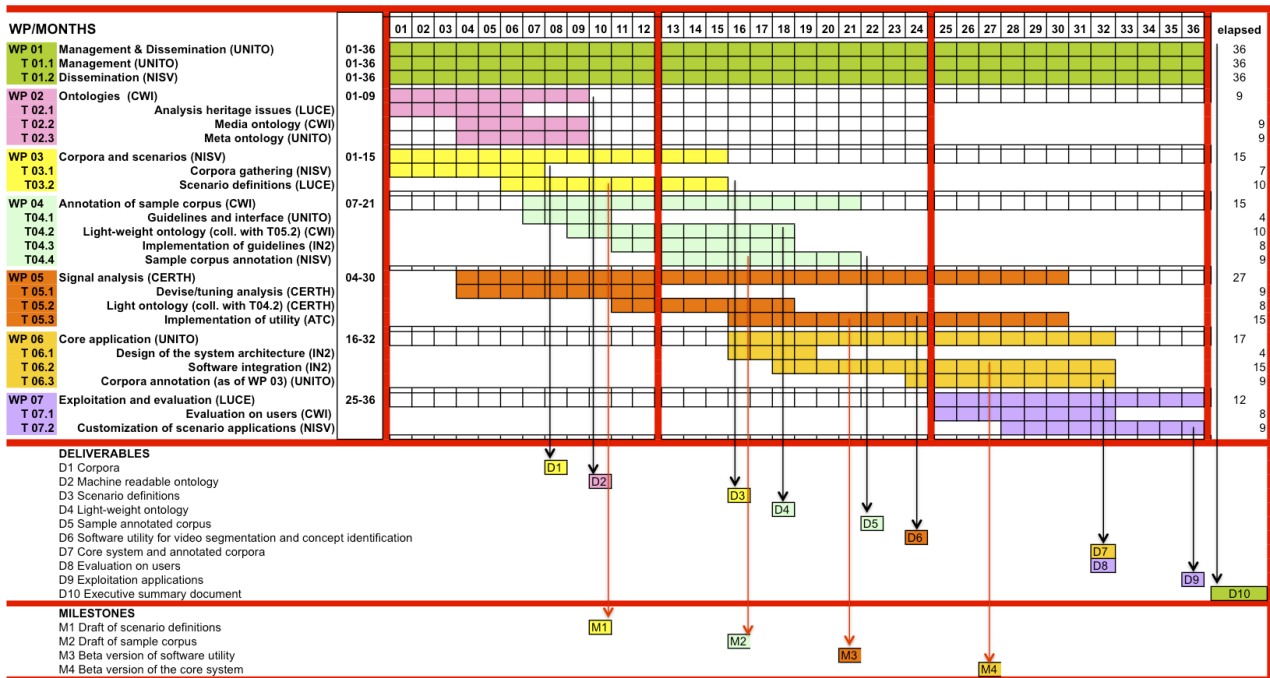
While these activities are proceeding, another workpackage (WP 5) works on devising algorithms and techniques for the analysis of the low-level features of the audiovisual signal: in particular, we work on techniques for the video segmentation (or scene boundary detection) and for the concept detectors. The latter will be tuned to the light-scale ontology, working in tight collaboration with the ontology expert partners. All this work will deliver a software utility that will contribute to automatize the annotation of clips and their role in the structure of a complex drama media object. This workpackage will be led by the signal analysis partner of CERTH, in collaboration with the industrial partner ATC, and the contribution of the ontology partners and the content partners.

When both annotation interface and automatic utilities are on the way to be released, we devise a system architecture that can be at the core of the exploitation applications. Such architecture integrates the interfaces for the annotation and the software utilities for the automatic analysis, that will be customized for the exploitation scenarios. The system will provide an interface for being customized on the scenario application with little implementation effort through a sort of parameter selection for driving the search operation. Finally we will apply the system to the annotation of the corpora gathered by the content institutions, and that will be at the core of the exploitation applications. This workpackage (WP 6) is led by UNITO, with the strong contribution of the industrial partners IN2 and ATC on the design and implementation, and the participation of the content institutions in assisting the annotation phase.

The last workpackage (WP 7) concerns the realization of the exploitation applications, through mostly parameter setting on the core system. Content institutions will identify the issues that need to be addressed and will customize the exposed parameters, possibly with the intervention of the industrial partners. The academic partners will conduct the evaluation studies on system users to assess the validity of the approach and to report on possible redesigns and optimizations.

ii)   **Show the timing of the different WPs and their components (Gantt chart or similar).**

We adopt an approach to timing that interprets a strong collaboration among WP's and single tasks. This is reflected in the Gantt below, with several overlaps over WP's and tasks for time elapses, and later in the management procedure, with technical meetings where people share their ideas from multidisciplinary viewpoints.

WP/MONTHS		elapsed
WP 01 Management & Dissemination (UNITO)	01-36	36
T 01.1 Management (UNITO)	01-36	36
T 01.2 Dissemination (NISV)	01-36	36
WP 02 Ontologies (CWI)	01-09	9
T 02.1 Analysis heritage issues (LUCE)		
T 02.2 Media ontology (CWI)		9
T 02.3 Meta ontology (UNITO)		9
WP 03 Corpora and scenarios (NISV)	01-15	15
T 03.1 Corpora gathering (NISV)		7
T03.2 Scenario definitions (LUCE)		10
WP 04 Annotation of sample corpus (CWI)	07-21	15
T04.1 Guidelines and interface (UNITO)		4
T04.2 Light-weight ontology (coll. with T05.2) (CWI)		10
T04.3 Implementation of guidelines (IN2)		8
T04.4 Sample corpus annotation (NISV)		9
WP 05 Signal analysis (CERTH)	04-30	27
T 05.1 Devise/tuning analysis (CERTH)		9
T 05.2 Light ontology (coll. with T04.2) (CERTH)		8
T 05.3 Implementation of utility (ATC)		15
WP 06 Core application (UNITO)	16-32	17
T 06.1 Design of the system architecture (IN2)		4
T 06.2 Software integration (IN2)		15
T 06.3 Corpora annotation (as of WP 03) (UNITO)		9
WP 07 Exploitation and evaluation (LUCE)	25-36	12
T 07.1 Evaluation on users (CWI)		8
T 07.2 Customization of scenario applications (NISV)		9

DELIVERABLES
D1 Corpora
D2 Machine readable ontology
D3 Scenario definitions
D4 Light-weight ontology
D5 Sample annotated corpus
D6 Software utility for video segmentation and concept identification
D7 Core system and annotated corpora
D8 Evaluation on users
D9 Exploitation applications
D10 Executive summary document

MILESTONES
M1 Draft of scenario definitions
M2 Draft of sample corpus
M3 Beta version of software utility
M4 Beta version of the core system

In the Gantt chart above, the top row reports the month number and the several WP's are marked with a different color. For each WP, we report the time span of each task and the total time span of the WP (all with the same color). Below, Deliverables and Milestones are reported, at the month of delivery. There are ten deliverables and milestones; deliverables and milestones are colored with the WP color that is responsible for that delivery. The rightmost column reports the time elapses for each WP and task respectively.

iii)   **Provide a detailed work description broken down into work packages:**
- **Work package list (please use table 1.3a);**
- **Deliverables list (please use table 1.3b);**
- **List of milestones (please use table 1.3c)**
- **Description of each work package (please use table 1.3d)**
- **Summary effort table (1.3e)**

Table 1.3 a: Template - Work package list

Work package list

Work package No	Work package title	Type of activity	Lead partic no.	Lead partic. short name	Person-months	Startm onth	Endmo nth[5]
1	Management and Dissemination	Management	1	UNITO	31	1	36
2	Definition and tuning of ontologies	RTD	2	CWI	40	1	9
3	Corpora and scenarios	RTD	6	NISV	51	1	15
4	Annotation of sample corpus	RTD	2	CWI	68	7	21
5	Automatic analysis of video signal	RTD	3	CERTH	76	4	30
6	Core application for annotation and search	RTD	1	UNITO	62	16	32
7	Exploitation and evaluation	DEMO	7	LUCE	58	25	36
	TOTAL				386		

Table 1.3 b: Template - Deliverables List

List of Deliverables

Del. no.	Deliverable name	WP no.	Nature	Dissemi-nation level	Delivery date (proj. month)
1	D1 Corpora	3	O	RE	8
2	D2 Machine readable ontology	2	P	RE	10
3	D3 Scenario definitions	3	R	PU	16
4	D4 Light-weight ontology	4	P	RE	18
5	D5 Sample annotated corpus	4	D	PU	22
6	D6 Software utility for video segmentation and concept identification	5	P	RE	24
7	D7 Core system and annotated corpora	6	P	RE	33
8	D8 Evaluation on users	7	R	PU	33
9	D9 Exploitation applications	7	D	PU	36
10	D10 Executive summary document	1	R	PU	36

Table 1.3c Template - List of milestones

Milestones

Milestones are control points where decisions are needed with regard to the next stage of the project. For example, a milestone may occur when a major result has been achieved, if its successful attainment is a required for the next phase of work. Another example would be a point when the consortium must decide which of several technologies to adopt for further development.

Milestone number	Milestone name	Work package(s) involved	Expected date	Means of verification
1	M1 Draft of scenario definitions	3	10	assessed and approved by stakeholders in each scenario
2	M2 Draft of sample corpus	4	16	expert verification about diversification and coverage
3	M3 Beta version of software utility	5	21	released and running on a subset of platforms
4	M4 Beta version of the core system	6	27	debugged and running on target platforms

Table 1.3 d: Template - Work package description

Work package description

WP 1 - Management and dissemination

WP leader: UNITO

Work package number	1		Start date or starting event:		Kick off		
Work package title	**Management and dissemination**						
Activity type	Management 100%						
Participant number	7						
Participant short name	UNITO	CWI	CERTH	IN2	ATC	NISV	LUCE
Person-months per participant	12	2	4	3	2	4	4

Objectives

The goal of WP1 is to manage the overall project by: starting the workpackage subgroups, checking the work advancements, identifying possible critical points and triggering alternative solutions, assessing the partial and final results on the technological pipeline, assessing the results on the annotation of cinematographic corpora and the exploitation applications. Given the interdisciplinary nature of the project, one of the objectives of the Management workpackage is the sharing of constraints and methods coming from the various partners. Finally, the WP1 checks the delivery of the deliverables and the achievement of milestones.

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

The workpackage WP1 consists of two tasks, Management and Dissemination, respectively. All the partners will contribute to this workpackage: UNITO and CWI will lead the two tasks; all the other partners contribute to management and dissemination, especially those who lead other workpackages (CERTH, NISV and LUCE).

T 1.1. Management (task leader: UNITO, months 1-36)

Since the project is carried out through a thorough integration of academic, industrial and content partners, the management needs to favor an effective implementation of knowledge transfer and objective sharing among partners:
- from the academic partners: ontology knowledge and formalization of media object representation to the industrial partners and the content institutions;
- from content institutions: targeted domains for light-scale ontologies and scenario definitions to academic partners and industrial partners;
- from industrial partners: request of detail for models to academic partners and constraints on the feasibility of scenario applications to content institutions.

These communications and transfers will be achieved through the organization of seminars and workshops at the startup of the activities (in conjunction with technical meetings) in order to inform one another about the respective competences and to set up a repository for the selected documentation and a wiki for the retrieval of solutions on specific problems that come out during the project.

Then, during the advancement of the project, the management will follow the achievement of results: we have set four milestones in order to verify the incomplete, though advanced releases of some crucial software module or document/corpus, namely the drafts of the scenario definitions and of the annotated sample corpus and the beta versions of the video analysis software utility and the core annotation/search system.

The task T 1.1 will also address the possible risks that can come out from critical points, through a specific monitoring of the activities (as reported below, at point V of this section) and the organization of specific tests for each technology that is included in the system implementation.

Meetings will be organized for the achievement of milestones and the presentation of major deliverable to be shared among partners, when possible in conjunction with the dissemination activities (see below).

T 1.2. Dissemination (task leader: CWI, months 1-36)

Given the interdisciplinary nature of the project, great importance will be given to disseminate the work of the project in journals, conferences, and forums of different areas: cinematic video analysis (computer vision/multimedia communities), cinematic metadata (linked open data/semantic web community); access, learning, and re-purposing of cinematographic heritage (cultural heritage community); computational storytelling and multimedia production (interactive storytelling community).
During the course of the project we will organise three technical workshops, to facilitate: knowledge sharing in the communities involved with the project; project promotion; results dissemination and consolidation of the technologies used in the project. The workshops will be held in the areas listed above.
Towards the end of the project, we will organise a fourth workshop oriented around the annotation, studying and re-purposing of film footage, with a cross-set of attendees from content providers (user community) and computer scientists (developer community). The aims of this event are to: promote the results of the project at a European or international cultural heritage event and act as input to future EU policy on exploitation of cinematographic heritage (as claimed by reports at the EU level, see e.g., http://www.dae-filmheritage.eu/). Finally, we will open the access to the application to a community of users that can try the project results to annotate and search on some audiovisual object; the users can also comment on the application and provide precious tests for design purposes and user studies.

The use of web-base tools will facilitate and enable the communication and the cooperation among the partners. For institutional documentation, a web based repository will allow the partners to be aware of the advancement of administrative processes and remain updated on the scientific advancement of the project. Moreover, social tools will enable the sharing of resources and the free discussion, improving the cohesion of the research efforts at the consortium level. At the same time, the web site will host and facilitate the conduction of the dissemination activities, informing the scientific community about the project's advancement and results and involving the community of the potential users in the project conduction since its initial phases.

Deliverables (brief description) and month of delivery

D 10: **Executive summary document on "The role of semantic annotation in the exploitation of the cinematographic heritage" delivered at month 36**. In this report, whose writing will be guided by the content institutions NISV and LUCE, the experience of the project CACHE-22 will be described thoroughly to benefit other content institutions around Europe for the exploitation of their archives.

Work package description

WP 2 - Specification and tuning of project ontologies

WP leader: CWI

Work package number	2		Start date or starting event:		Kick off	
Work package title	Specification and tuning of project ontologies					
Activity type	RTD 100%					
Participant number	5					
Participant short name	UNITO	CWI	CERTH	IN2	ATC	NISV	LUCE
Person-months per participant	15	18	2	0	0	1	4

Objectives

Two ontologies that define the content form of the metadata that will annotate the media objects:

- a media ontology, that incorporates the features of media objects in general, including the Dublin Core metadata for cultural resources;

- a drama ontology, that introduces an internal structure for the organization of the narrative scenes of a dramatic media object and the actional content of each scene.

The two ontologies must take into account the specific issues of cinematographic heritage.

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

The WP2 will work by analyzing the issues carried on by the cinematographic heritage on the ontological format; then, we will work on the two ontologies, for media and dramatic features, respectively, that is the structure of the media objects (e.g., the field "Media item creator") and of dramatic elements (e.g., the field "Location of the action"). This workpackage (WP2) is based on previous experiences of two of the partners, CWI and UNITO, respectively. In particular, CWI has developed a specific ontology to describe media, COMM. COMM is a Core Ontology for MultiMedia that allows attaching concepts from an existing Semantic Web ontology to any arbitrary description of part of a media asset. For what concerns media description, W3C is preparing a media fragment recommendation, for specifying a spatial/temporal fragment of a media asset (http://www.w3.org/TR/media-frags/). Since media fragments are identified by a URI, they can be used in RDF, according to the linked data paradigm. UNITO has developed a generic ontology for the narrative features, called Drammar, in the CADMOS project (www.cadmos-project.org); finally, the content institutions, led by LUCE, will contribute with the analysis of the issues that are necessary for tuning the ontologies on audiovisual corpora, as well as defining new concepts. In WP2 we also agree on the common tools for representing the ontologies, as well as the reasoning system that will be employed in the project. The WP2 consists of three tasks.

T 2.1. Analysis of cinematographic heritage issues (task leader: LUCE, months 1-4)
In this task, in collaboration with 3.1 (on the gathering of corpora), the partners will analyze the issues carried on by the objects contained in the cinematographic heritage archives and the impact of such issues for the ontological encoding. The issues can concern both elements that are intrinsic of the objects (e.g., aspect ratio of the film) and elements that are useful for the exploitation of the objects (e.g., a specific location where the object was filmed).

T 2.2. Tuning of the media ontology for narrative objects (task leader: CWI, months 4-9)
In this task, working on the results of the task 2.1, namely the specific issues concerning the encoding of the items of the cinematographic heritage, the partners will identify and adapt the COMM media ontology for the encoding. The result is the compilation of a media ontology in a machine readable format, as agreed with the partners of the task 3.2 (scenario definition).

T 2.3. Tuning of the ontology for narrative features (task leader: UNITO, months 4-9)

In this task, working on the results of the task 2.1, namely the specific issues concerning the encoding of the dramatic features of a media object, the task will build upon the Drammar ontology (elaborated by UNITO for previous projects on drama production). In particular, we will take into account the issues connected to the exploitation in terms of learning and access of cultural heritage, so inserting ontology elements that were missing, together with checking the functioning of reasoning services. The ontology, in a machine readable format, shall satisfy the requirements of the task 3.2 (scenario definition), with respect to the applications that will be built later in the project. Also, the tasks 2.2 and 2.3 must identify some common encoding, in order to create a unique ontology format.

Deliverables (brief description) and month of delivery

D2. A **machine readable ontology**, that incorporates the features of media objects in general and the dramatic features of media objects, **with associated documentation**. **Delivered at month 9**.

Work package description

WP 3 - Corpora and scenarios

WP leader: NISV

Work package number	3			Start date or starting event:		Kick off	
Work package title	**Corpora and scenarios**						
Activity type	RTD 100%						
Participant number	7						
Participant short name	UNITO	CWI	CERTH	IN2	ATC	NISV	LUCE
Person-months per participant	15	8	2	2	2	10	12

Objectives

1. Definition and gathering of the corpora that will work as a test for the technology pipeline

2. Setting of the use case scenarios and application design

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

The workpackage WP3 addresses the definition of the corpora that will work as a test for the technology pipeline, and the setting of use case scenarios concerning the same corpora, that will inspire the design of exploitation applications for access and learning of cultural heritage. Corpora will be used by the WP's 2, 4 and 5 for the definition of the ontologies and the light-scale content ontology; use case scenarios will be used by WP's 6 and 7 for the definition of the core annotation/search system and the exploitation applications. We also extract from the corpora, a representative sample corpus for immediate testing of the technologies. WP 3 is mostly carried out by the two content partners, namely NISV and LUCE, with the contributions of the academic researchers and the industrial partners. WP3 consists of two tasks.

T 3.1. Proposal and gathering of corpora (task leader: NISV, months 1-7)

In this task the content institution will identify the corpora that will be the application domain for the exploitation (WP7). The content institutions will base such a choice on the recent appraisals from the generic audience about the cinematographic heritage, teaching programmes in academies, schools and universities, the indications from the EU Commission on the cultural heritage dissemination, their current internal interests as archives and stock companies. These motivations will be negotiated with the constraints posed by the technological pipeline, namely the current possibilities in ontological representations, video playing platforms, web based reasoning devices, effectiveness of exploitation applications. This task may also include digitalization of some items. From the corpora, we will also extract a sample for testing and debugging the beta versions of the applications to be implemented in the subsequent workpackages.

T 3.2. Scenario definition for cultural heritage access and learning (task leader: LUCE, months 4-9)

This task will envision the possible scenarios of exploitation for the cinematographic heritage, and in particular for the corpora identified in the task 3.1. The same motivations that were applied to the task 3.1 for the selection of corpora holds here, negotiated with the constraints posed by the application technology and the hypothesized target users. The result is a technical document illustrating the scenarios, with indications on the interfaces, the workflow needed to operate in the scenario, possible constraints on technologies. This work will account for the inputs and requirements collected by surveying the needs expressed by the stakeholders in education, and media education in particular.

Deliverables (brief description) and month of delivery

D1. Machine readable corpora of cinematographic heritage, with a sample corpus used in testing, and **associated index and documentation**. **Delivered at month 8**.

D3. A document illustrating the use case scenarios and the application design, with indications on the interfaces, a draft architecture, the workflow needed to operate in the scenario, technologies involved. **Delivered at Month 16**.

Work package description

WP 4 - Annotation of sample corpus

WP leader: CWI

Work package number	4		Start date or starting event:		Month 7	
Work package title	Annotation of sample corpus					
Activity type	RTD 100%					
Participant number	6					
Participant short name	UNITO	CWI	CERTH	IN2	ATC	NISV	LUCE
Person-months per participant	24	24	2	8	0	4	6

Objectives

- Annotation guidelines

- Interface for manual annotation

- Light-scale commonsense ontology

- Sample annotated corpus

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

Based on the WPs 2 and 3, this WP works at the design of annotation guidelines and a suitable interface for realizing a sample annotated corpus of dramatic media objects. It also elaborates a light-scale commonsense ontology, driven by the actual content selected for the exploitation and in collaboration with the partners working on the signal analysis. The light-scale ontology is useful for providing annotation terms that can be shared among applications. The light-scale ontology will be built in collaboration with the WP5 on the automatic analysis of video, in particular with the task on the definition of concept detectors. WP4 will produce the sample annotated corpus that will be the golden standard for inspiration and evaluation of automatic phases. WP4 consists of four tasks. The conduction of all the tasks will account for the annotation practices and metadata standards of related initiatives, especially at the European level, such as Europeana.

T 4.1. Guidelines and interface design for manual annotation of narrative media objects (task leader: UNITO, months 7-10)

Task 4.1 will define the guidelines and the interface for creating a sample annotated corpus. Based on the corpora identified and the use case scenarios, the task will lay down the guidelines for annotating the sample corpus and will design an interface for the manual annotation of the corpus. This work will be led by UNITO, based on its experience of the project CADMOS (that involved a manual annotation of narrative audiovisuals), with the support of the content istitutions and IN2, that will implement the interface to the ontologies, that record the annotated media instances. The result is documentation on the guidelines for the annotators, the interface and the interaction design.

T 4.2. Elaboration of light-scale ontology (collaboration with T 5.2) (task leader: CWI, months 9-16)

Task 4.2 will build a light-scale ontology, based on the corpora defined previously and the exploitation scenarios. Following the paradigm of linked data, the task proceeds by identifying concepts in the corpora (starting from the sample corpus) and introducing them as URI's in commonsense ontologies that can be identified (such as, e.g., YAGO-SUMO). The task will balance between completeness of descriptions (in order to be able to carry out reasoning tasks) and effectiveness in being the reference for the automatic analysis of video.

T 4.3. Implementation of guidelines through the designed interface (task leader: IN2, months 11-18)

Task 4.3 will develop the interface for the ontologies defined in WP 2 in order to allow the human annotation of the sample corpus. The interface will implement the guidelines for annotation defined in the task 4.1 and will refer to the light-scale ontology defined in the task 4.2. Through the annotation interface, the human experts to annotate the sample corpus.

T 4.4. Creation of golden standard annotated sample corpus (task leader: NISV, months 13-21)

This task will accomplish the annotation of the sample corpus. Content institutions will work under the guidance of research institutions, that will provide training and assistance in the annotation phase. The annotators will be historians of cinema or narratologists, with a training in logic and ontologies, provided by the research institutions. The annotation will proceed with at least two independent annotations, supervised by a senior expert from the content institutions, and will mediate between expressive description of content and effectiveness in providing terms for later search.

This task includes a **Milestone, M2**, at month 16, that is a draft release of the annotated sample corpus, to check the validity of the annotation approach, before the actual release of the golden standard corpus at month 22.

Deliverables (brief description) and month of delivery

D4. Light-weight commonsense ontology, derived from a collection of large scale ontologies, in a machine readable format. **Delivered at Month 18**.

D5. Sample annotated corpus of audiovisual objects, described according to the media ontology and the drama ontology and annotated with the concepts from the light scale ontology, with the associated documentation about the audiovisual objects (expressed according to the appropriate metadata standards) and the informal documentation about the annotation process . **Delivered at Month 22**.

Work package description

WP 5 - Automatic analysis of video signal

WP leader: CERTH

Work package number	5		Start date or starting event:		Month 4	
Work package title	**Automatic analysis of video signal**					
Activity type	RTD 100%					
Participant number	6					
Participant short name	UNITO	CWI	CERTH	IN2	ATC	NISV	LUCE
Person-months per participant	5	0	35	3	24	3	6

Objectives

Software utilities for

1. video segmentation into meaningful scenes

2. concept identification within scenes

through the analysis of the low-level features of the audiovisual signal and the connection with the light-scale ontology.

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

The workpackage 5 works on devising algorithms and techniques for the analysis of the low-level features of the audiovisual signal: in particular, WP5 works on techniques for the video segmentation (or scene boundary detection) and for the identification of concepts from concept detectors. The latter will be tuned to the light-scale ontology, working in tight collaboration with the task 4.2. This delivers a software utility that will contribute to automatize the annotation of clips and their role in the structure of a complex drama media object. This workpackage will be led by the signal analysis partner of CERTH, in collaboration with the industrial partner ATC, and the contribution of the ontology partners and the content partners.

T5.1 Video fragmentation (task leader: CERTH, months 4-12)

This task will deal with the decomposition of video to meaningful fragments, which will serve as the elementary pieces of audio-visual content that will be accessed, annotated and re-used. In this context, meaningful fragments are not only the video shots and scenes, but also sub-shots that depict a single action in isolation from other actions that may be part of the same shot. For the purpose of shot and scene segmentation, state of the art techniques will be adapted to cinematographic content, taking into account information that is central to the nature of this content, such as actor occurrences. The latter will be derived by means of automatic face/person detection and unsupervised face clustering, so that different instances of the same person across the video are grouped together and are used for assessing the possible video decompositions. For sub-shot detection, techniques that have been developed for shot and scene segmentation (e.g., transition graphs) as well as for action segmentation in video (e.g., motion-based techniques) will be combined and adapted to the tasks of fragmenting a shot to its constituent sub-shots.

T5.2 Light ontology and concept detectors (CERTH, months 11-18)

In this task a light ontology will be identified, and the corresponding detectors will be developed. The ontology definition will be performed in collaboration with WP4, so as to reconcile the needs of annotation, in the context of cinematographic applications, with the capabilities of the state of the art concept- and event-detectors that will be developed in this task. Narrative features (ontological concepts) will refer to characters, objects, environments (locations) involved, as well as the actions carried out by the characters; some concept detectors can directly fill up some concepts (for example, a message conveyed verbally will be annotated as

some content that is the result of a speech-to-text conversion).
This light ontology will include basic relations that can be exploited by automatic video analysis, e.g. parent-child relations, disjunctions, etc. The detectors to be developed will allow associating video fragments with the concepts of the defined ontology. They will leverage ongoing work in the context of the TRECVID benchmarking activity, and extend the current state of the art both in terms of effectiveness (i.e., how accurately the concepts are detected) and in terms of computational efficiency. The efforts towards these goals will be two-fold. On the one hand, techniques for video description based on keyframe extraction and analysis will be extended to the analysis of other 2D visual information structures that also encompass time information, without increasing the computational complexity of the concept detection process. For this, the notion of video tomography will be investigated. On the other hand, the possibility of using ensembles of (very fast) linear classifiers, instead of (much slower) individual kernel-based classifiers, will be investigated. The emphasis there will be on very efficiently exploiting a large number of initially high-dimensional low-level features, instead of performing more elaborate but much more computationally demanding kernel-based learning.

T5.3 Implementation and evaluation on the sample corpus (ATC, months 16-30)

The techniques developed in tasks T5.1 and T5.2 will be implemented here as a set of robust software utilities that can be used for processing large volumes of video content. For achieving this, implementation-specific software optimization will be undertaken in this task (where the goal will be to accelerate the software implementation of the key analysis algorithms by modifying its implementation, without making changes to the underlying mathematical formulation), complementing the algorithm-specific efforts towards minimizing the algorithms' computational complexity that are an inherent part of the work in T5.1 and T.5.2. The resulting software utilities will be tested and evaluated, in terms of effectiveness and computational efficiency, on the sample corpus used throughout the project, as well as on suitable external datasets.

At month 21 we have inserted a milestone, M3, concerning a beta release of the software utility.

After releasing the software utility at month 24, CERTH and ATC will provide tuning and bug fixing for the use of the software utility within the core system, developed in the WP 6. This task will be in collaboration with the tasks 6.2 and 6.3.

Deliverables (brief description) and month of delivery

D6. Software utility for video segmentation and concept identification, accompanied by the documentation (UML diagrams, API reference) and a user guide with examples and code snippets.
Delivered at Month 24, with the Milestone **M3** on the beta version.

Work package description
WP 6 - Core application for annotation and search
WP leader: UNITO

Work package number	6		Start date or starting event:	Month 16		
Work package title	Core application for annotation and search					
Activity type	RTD 100%					
Participant number	5					
Participant short name	UNITO	CWI	CERTH	IN2	ATC	NISV	LUCE
Person-months per participant	24	15	3	15	5	0	0

Objectives

Core system for annotation and search of dramatic media objects

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

WP 6 devise a system for annotation and search of dramatic media object, that will be at the core of the exploitation applications. Such an architecture integrates the interfaces for the annotation and the software utilities for the automatic analysis, that will be customized for the exploitation scenarios with little implementation effort through a sort of parameter selection for driving the search operations. We will apply the system to the annotation of the corpora gathered by the content institutions in WP3, and that will be the object of the exploitation applications. WP 6 is led by UNITO, with the strong contribution of the industrial partner IN2 on the system realization, and the support CERTH and ATC on the design and implementation, and the participation of the content institutions in assisting the annotation phase. WP 6 consists of three tasks.

T 6.1   Design of the system architecture (IN2 , months 16-19)

Following the outcome of the scenario definition (T3.2), this task aims to draw up the system architecture. The approach we will follow starts by translating the user requirements into functional requirements, grouping functional requirements into components, joining components into modules and architecture layers and finally describing the architecture. This will allow us to present different level of details from a high-level block diagram to a detailed view of the functionalities and how they interact together. Interfaces for communication and interaction between those components will also be defined. As also suggested in "http://www.dae-filmheritage.eu/mediapool/100/1000452/data/exec_summary_en.pdf", we aim at a for the application to be web based, so as to grant online access to all the stakeholders in film and media community and to the general public.

T 6.2   Implementation of the integrated software (IN2, months 18-32)

This task will be in collaboration with the task 5.4, for the bug fixing of the automatic procedures of video analysis. Moreover, it will also keep the software open to the integration of existing applications (for metadata editing, semantic browsing, production management, etc.) as plugins to the system architecture. The sharing of annotation on social networks and media sharing platforms will also be taken into account.

T 6.3   Annotation of the corpora identified in WP 03 (UNITO, months 24-32)

In this task, we will annotate the corpora identified in the task 3.2. Taking the deliverable D 4.2, the sample annotated corpus, as a reference, we apply the core system in the proposal of automatic segmentations and

concept annotation. These annotations can be approved or changed by the annotators. Also this task will in collaboration with the task 5.4, because of the possible tuning of the automatic procedures, as the annotation proceeds.

At the month 27 we have a **milestone**, **M4**, for the beta release of the core system, together with an annotated corpus sample (to be compared with the golden corpus delivered in WP 4), that concerns both tasks 6.2 and 6.3.

Deliverables (brief description) and month of delivery

D7. Core system for annotation and search of dramatic media objects and annotated corpora, with associated documentation, and a user manual describing the use of the system through practical examples, taken from different scenarios and corpora. **Delivered at month 32**, with Milestone M4 for the beta version of the system and the annotation of the sample corpus with the core system.

Work package description
WP 7 - Evaluation and exploitation
WP leader: LUCE

Work package number	7		Start date or starting event:		Month 25	
Work package title	**Exploitation and evaluation**					
Activity type	DEMO 100%					
Participant number	7					
Participant short name	UNITO	CWI	CERTH	IN2	ATC	NISV	LUCE
Person-months per participant	15	12	2	4	2	8	15

Objectives

1. realization of the exploitation applications

2. evaluation studies on system users

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

The last workpackage (WP 7) concerns the realization of the exploitation applications, mostly accomplished through parameter setting on the core system. The academic partners will conduct the evaluation studies on the core system users to assess the validity of the approach and to report on possible redesigns and optimizations. Content institutions will identify the issues that need to be addressed and will customize the configuration parameters, possibly with the intervention of the industrial partners.

T 7.1   Evaluation of the core system on users (CWI, months 25-32)

The evaluation of the core system will be conducted by testing it with real users, selected in fields identified in the scenario definition (cultural heritage institutions, educational institutions, etc..) in WP 3. Mainly inspired by the paradigm of user studies, the evaluation requires, for each scenario, the specification of the following elements: the functionalities to be tested (the objective of the evaluation), the user tasks by which the evaluation objectives will be reached, and the specification of the statistical and analytical tools that will be employed to measure the system performance with the users. The satisfaction of the users will also be measured by means of questionnaires and depth interviews, whose results will be integrated in a final report.

T 7.2   Customization and testing of scenario applications (NISV, months)

Given the results of the previous task, the annotation and search system will be customized to fit the needs of specific use scenarios. Different system configurations will be applied to different scenarios, and their appropriateness for executing the tasks required by the scenario will be evaluated with the help of experts. The cost and time required for the customization will also be assessed, with the goal of gathering data for the possible exploitation of the project results.

Deliverables (brief description) and month of delivery

D8. Evaluation report. Delivered at month 32, D8 describes the experimental setting designed for the evaluations on the core system of annotation and search delivered in WP 7, and its methodological

assumptions, presenting and discussing the evaluation results and the possible redesign issues and customization possibilities for the task 7.2.

D9. Exploitation applications, a set of different configurations of the core systems, designed and tested on the specific scenarios, accompanied by descriptive use cases. **Delivered at month 36.**

*Table 1.3e   Summary of effort*

Summary of effort

A summary of the effort is useful for the evaluators. Please indicate in the table number of person months over the whole duration of the planned work, for each work package by each participant. Identify the work-package leader for each WP by showing the relevant person-month figure **in bold**.

Partic. no.	Partic. short name	WP1	WP2	WP3	WP4	WP5	WP6	WP7	Total person months
1	**UNITO**	**12**	15	15	24	5	**24**	15	**110**
2	**CWI**	2	**18**	8	**24**	0	15	12	**79**
3	**CERTH**	4	2	2	2	**35**	3	2	**50**
4	**IN2**	3	0	2	8	3	15	4	**35**
5	**ATC**	2	0	2	0	24	5	2	**35**
6	**NISV**	4	1	**10**	4	3	0	8	**30**
7	**LUCE**	4	4	12	6	6	0	**15**	**47**
Total		**31**	**40**	**51**	**68**	**76**	**62**	**58**	**386**

iv)   Provide a graphical presentation of the components showing their interdependencies (Pert diagram or similar)

The PERT diagram below reports all the tasks in rectangular shape (only WP1, on top, does not distinguish on tasks), with their interdependencies. For color prints, tasks are colored with the same color used in the Gantt chart. For each task, we indicate the time elapse and a short version of the title. All deliverables (white circles), indicated by the label, short title, and the delivery month in parenthesis, are connected with the tasks that produce them and are input to tasks that use them. To four tasks (T3.2, T4.4, T5.3, T6.2), we have attached boxes (in red/dark, with white characters), indicating the milestone involved, again indicating label, month, and short title. The tasks T4.2 and T5.2 are joined in double-colored single box, to represent the intense collaboration of the two tasks (they deliver a single deliverable D4).
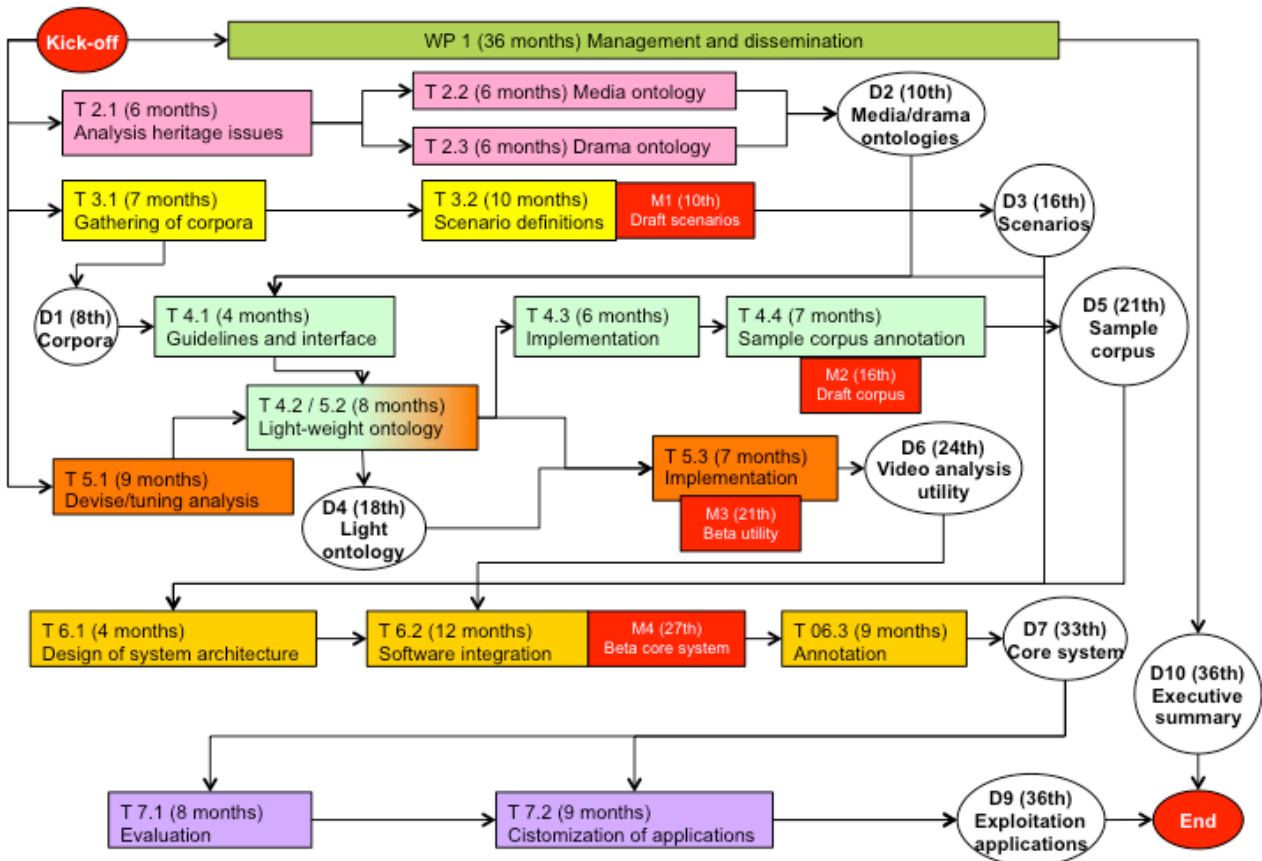
v) Describe any significant risks, and associated contingency plans

There can be some risks in the achievement of the project results.

Ontological concepts that are tricky to identify automatically
The most risky achievement is the devise of automatic procedures for filling the semantic gap at a high level, that is for connecting automatically the concept detectors and the ontological concepts in some tricky cases. For example, although there have been some results in video surveillance and sport news applications, the general problem of action/event recognition is still far from being solved (see e.g. Detecting Video Events Based on Action Recognition in Complex Scenes Using Spatio-Temporal Descriptor, Zhu et al. at ACM MM 2009). Since events and actions are part of the description of the contents of a narrative media object, project CACHE-22 will try to avoid tricky situations by including the partners that deal with the automatic procedures into the WP's concerning the selection of the corpora to be used in the scenarios.
However, in case of problematic situations we resort to a contingency plan that involves manual annotation, by taking into account some commonsense ontology, or limiting annotations to object/character involved in the action.

Elaboration of the light-scale ontology
The elaboration of the light-scale ontology can present difficulties because of divergent issues to be represented. This depends on the audiovisual content of the items that enter the corpora selection: for example, if the light-scale ontology must cover a large portion of a commonsense knowledge base to account for the domains represented in the corpora, then the goal of keeping the ontology manageable fails, with consequent troubles on the effectiveness of annotation as well as on concept detectors. This constraint of keeping light scale the ontology will be enforced again in the task of corpora selection (3.1) as well as during the compilation of the ontology (4.2).
However, in case of troubled situations we can resort to insert specific concepts under some default category or, equivalently, to add a free, limited list of concepts.

Web based platform for annotation and scenario applications
The design of effective search and retrieval functionalities for the exploitation of audiovisual archives is also a risk factor. The multiplicity of target audiences, ranging from scholars and eductions stakeholders to the general public, also puts at risk the design of effective search and retrieval tools. In order to minimize this risk, the project will define scenarios, along with content partners, that will inform the development of the core application and of the end-user scenarios; also, we will adopt a user-centric perspective in the development of the end-user applications.
As a recovery plan, we envisage the integration, along with the institutional metadata and annotation, of user-generated tags and description that will enrich and accompany the meaning provided by the narrative model underlying the annotation, in a Web 3.0 perspective.

Section 2. Implementation
(Recommended length for Section 2.1 - five pages)

2.1 Management structure and procedures
Describe the organisational structure and decision-making mechanisms of the project. Show how they are matched to the complexity and scale of the project.

CACHE-22 will implement a lightweight management procedure. The management of the project is indented to be light and relying on horizontal network rather then on strict vertical hierarchy.

Management Board
The project will be led by the Management Board (MB), which will consist of the 7 PIs, one from each of UNITO (Lombardo), CWI (Hardman), CERTH (Kompatsiaris), LUCE (Rendina), NISV (Oomen), IN2 (Stan) and ATC (Sarris). The MB will be chaired by Lombardo.
The MB aims:
1. to review project progress and assess project performance;
2. to ensure the project maintains its technical objectives;
3. to ensure the project maintains its relevance within the specific program;
4. to resolve any technical, administrative or contractual issues where lower levels of the management structure have been unable to do so;
5. to maintain regular contact with the Work package Leaders in order to ensure effective progress of the project;
6. to prepare and enforce the project operational planning;
7. to review work progress;
8. to check the technical consistency between work packages and tasks;
9. to ensure that the overall project schedule is maintained through operational planning, and to anticipate and resolve deviations, whether from a technical or schedule standpoint;
10. to negotiate and confirm dates, venues and content for Technical Reviews;
11. to monitor the preparation, production and distribution of all deliverables;
12. to report project progress and to report any problems as they arise;
13. to ensure that all demands voiced in a technical review report are fully and promptly resolved.

Scientific Advisory Board
There will be a project Scientific Advisory Board (SAB). The purpose of the SAB is to review the project approaches, achievements and plans, and to help keep the project informed of developments in the wider context. We would select 3 candidates to take part in the scientific advisory board: one from the field of video processing, one from the field of cultural heritage content exploiter or narratologist, and one from the field of media communication.

Meetings
There will be a meeting of the MB (**MBM**) every 6 months, usually aligned with a technical meeting and/or a project milestone. All researchers on the project will be allowed to attend as well as the MB, in order to keep them engaged with the project and help with training for future leadership roles. Decision making will be by consensus; however, if a vote is necessary, each of the 7 PIs has a single vote.

Once a year the project will held a scientific project meeting (possibly aligned with MB meeting) in which the SAB (**SABM**) will be invited to discuss its view and consideration together with the other members of the research team.

There will be quarterly 1-2 day Technical Meetings (**TM)** to present and review research progress. Normally, representatives from all partners will attend and will be possibly aligned with MB meetings. In the
The TM takes in consideration all the technical and operational achievements, and problems, under the guidance of the Work Package Coordinators. Given the high interdisciplinary attitude of the CACHE-22

Project, the TCP will ensure a more effective, direct confrontation among project members, lowering the grade on jerarchyzation in the structure. Beside the general TM session, the coordinators can agree to launch sub session on meetings on specific topics with a restricted number of researchers and components. All the decisions taken by the TM will be under the consensus basis..

Role in the project

Project Coordinator
The Project Coordinator will oversee and co-ordinate all administrative aspects of the project. Its responsibilities in this respect are as follows:
1. to coordinate, collate and distribute all administrative and progress reports;
2. to maintain records of costs, resources and time-scales associated with the project;
3. to collect and submit to the Commission all individual cost statements of the partners;
4. to ensure prompt financial flows within the Consortium;
5. to remain in contact with the Commission regarding all contractual and administrative aspects of the project;
6. to maintain regular contact with all local administrative contacts;
7. to maintain regular contact and to meet regularly with the WPs to ensure effective progress of the project.

WP Coordinator
The WP Coordinator will oversee and co-ordinate all administrative aspects of the project. Its responsibilities in this respect are as follows:
1. to coordinate, collate and distribute all administrative and progress reports;
2. to maintain records of costs, resources and time-scales associated with the WP;
3. to stay weekly in contact with the Project Coordinator
4. to maintain overview the WP costs;
5. to maintain regular contact with the member of the WPand to meet regularly MB to ensure effective progress of the project.
6. to issue a technical review report to be submitted to the quarterly 1-2 day **technical meetings**

Services and supports
The project will set up listserve email lists for project wide day-to-day communications. In particular, two lists are initially proposed:
- an administrative issues list to communicate amongst the PIs,
- a general list to communicate amongst all project staff.
- single WP lists to communicate within the working team
-

A project web page will also record all project personnel along with contact email addresses.

A project wiki will be created to record ideas, design decisions, outstanding issues, etc. This will provide a collaborative noteboard for use inside the project. It is not intended for external communications.

Project documents and datasets, useful URLs and other information, will be made accessible through either the public website or internal wiki.

Information flow

The light quality of our management system will keep a constant flowing of information among the WP network via the service and support. Nevertheless the WP in charge of the management will make sure that the needed steps and passage will happen among every partner:
1. the Project Coordinator enforces the information flow between the MB members and the WP among the partners;
2. all the partners provide the Coordinator with their administrative information in order to facilitate the necessary relationship with the European Commission;

3. all the partners provide the WP Coordinator with the required information on resource management;
4. the WP Coordinator provide the required input and information to the partners involved in advancement of the work packages;
5. all documents produced in the project are made available to all partners. Project documents will be whenever possible distributed through electronic means;
6. free format communication and exchange between the various participants are encouraged.

Deliverables Handling

The project will have a Quality Assurance procedure. Each Deliverable document, dataset and software component will have an identified person from another project partner who will be responsible for reviewing the content or performance of the deliverable.

The MB will set up a project file for internal and external references. It will contain, beyond all approved technical documents, agendas, minutes of meetings and details of expenditure, the deliverables provided by the project.

Contract Management

Before contract signature, the partners agree to sign a consortium agreement that will govern their relations during the project.

2.2 Individual participants

(Maximum length for Section 2.2: one page per participant. However, where two or more departments within an organisation have quite distinct roles within the proposal, one page per department is acceptable.
The maximum length applying to a legal entity composed of several members, each of which is a separate legal entity, is one page per member, provided that the members have quite distinct roles within the proposal.)

For each participant in the proposed project, provide a brief description of the legal entity, the main tasks they have been attributed, and the previous experience relevant to those tasks. Provide also a short profile of the individuals who will be undertaking the work.

1. CIRMA - University of Torino (coordinator), contact person Vincenzo Lombardo,
vincenzo@di.unito.it
CIRMA (Centre for Research on Advanced Multimedia) is an interdisciplinary research centre of the University of Torino composed by two departments, Computer Science (Dipartimento di Informatica) and Human Studies (Dipartimento di Studi Umanistici), and 20 scholars, plus PhD students and Post-docs. CIRMA carries on and promotes research on the relationships between new and traditional media, with particular interest for artistic productions, practical applications and public installations. CIRMA participates into large and small scale projects at the local, national and international level, concerning computational storytelling, interactive installations, documentation models for contemporary art, formal representation for live performances, among which: *ReAl Code* (Intelligent encoding and search for media objects); 2010-12, *CADMOS* (Character-centred annotation of Dramatic Media Objects); the *Dramatour* project (dramatour.di.unito.it); the *Virtual Electronic Poem (VEP)* project, sponsored by the European Union – Culture 2000 programme, a VR reconstruction of the 1958 Poème électronique by Le Corbusier.

Role of UNITO in the project
On the basis of the local project CADMOS (http://cadmos-project.org), CIRMA has devised a computational framework for the semantic annotation of narrative audiovisual fragments; CIRMA has been developing a meta-ontology for the representation of the narrative features, the connection to a large scale ontology for the annotation of metadata, and a prototype corpus of semi-automatically annotated videos. Thanks to its interdisciplinary teamwork, CIRMA has been successful in putting the drama critics and the narratology studies into a computational perspective. The contributions to CACHE-22 will be: engineering of the meta-ontology, linguistic interface to a large-scale ontology, symbolic contribution low-level feature analysis, design of the narrative/dramatic features.

Key personnel
Vincenzo Lombardo PhD, Associate Professor of Informatics, University of Turin, Italy (www.di.unito.it/~vincenzo). He doctorated in Computer Science in the Turin-Milan University Consortium (1993). He is co-founder of CIRMA (Centre for Research on Advanced Multimedia). He is part of the Computer Science programme of the Doctoral School of Sciences and Innovative Technologies (http://dott-snti.campusnet.unito.it/do/home.pl). At the Virtual Reality & Multi Media Park public company (www.vrmmp.it), he runs the Art-Science Alliance Laboratory (ASA Lab - www.vrmmp.it/asalab/). He has been the coordinator of the EU-funded project Virtual Electronic Poem (www.edu.vrmmp.it/vep/) and of the national project DramaTour (www.dramatour.unito.it). Currently, he is the coordinator of the project CADMOS (www.cadmos-project.org), on topics strictly related to CACHE-22. His research concerns methodologies, models and applications of informatics to the production processes of multimedia artefacts and computational storytelling. His papers are published in international journals, books, conference proceedings. He carries on a production activity in multimedia communication, hosted by events at the international level.

Rossana Damiano PhD. Researcher in Computer Science at the University of Turin (http://www.di.unito.it/~rossana). She got her PhD in Computer Science from the University of Turin in February 2002, with a dissertation on "The role of norms in intelligent reactive agents". She is a member of the Interaction Models (Agents, Language and Expression) Group at the Department of Computer Science of the University of Turin, directed by prof. Leonardo Lesmo. She has been founder of CIRMA. She is leader in various research project (150 Digit, Labyrinth) an has played a core role in Dramatour and CADMOS. Her interests involve computational models of dialogue, socially intelligent agents, and interactive drama. Her papers are published in international journals, books, conference proceedings.

Antonio Pizzo PhD, Associate Professor of Drama and Performing Arts, University of Turin, Italy (www.cirma.unito.it/pizzo). He doctorated in Modern Theatre and Drama in the Naples-Salerno University Consortium (1997). He is co-founder (and present Director) of CIRMA. He is part of the Doctorate in Modern and Contemporary Theatre studies at Università degli Studi di Napoli "L'Orientale" (coord.: prof. Lorenzo Mango). He is cofounder and coordinator of Officine Sintetiche (www.officinesintetiche.it) a and is Director of the CRUD (a regional centre for contemporary dance) (www.crud.unito.it). He has collaborated with project DramaTour (www.dramatour.unito.it) and currently working in CADMOS (www.cadmos-project.org). His research concerns general theory of modern drama, model interactive drama and synthetic agents, digital performance, augmented stage and enhanced actors. He has published books in Italy and papers in international journals, conference proceedings.

2. Centrum Wiskunde & Informatica (CWI), contact person Lynda Hardman Lynda.Hardman@cwi.nl

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.

The most recent relevant related projects are: FP7 LinkedTV, where broadcast video assets are annotated and automatically linked and presented to end-user; FP7 Fish4Knowledge, where marine biologists are supported in accessing machine interpreted representations of large numbers of video images of fish; and the national SEALINCMedia project, where social networks are used to enrich access to media assets.

Role of CWI in the project
CWI has expertise in computational ontologies of media assets and integration with the linked open data paradigm. They will contribute experience in annotating video assets with discourse roles and using these to construct links among appropriate use cases. CWI will oversee the construction of the annotation application in the context of the media framework with the modules contributed by the other partners.

Key personnel

Prof. dr. **Lynda Hardman** is the head of the Information Systems department and part-time professor at the University of Amsterdam (http://homepages.cwi.nl/~lynda). 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 is a board member of Informatics Europe, and the Action Line Leader for ICT-mediated Human Activity in EIT ICT Labs.
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 VU University Amsterdam (http://homepages.cwi.nl/~jrvosse). He obtained his PhD from the VU University 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 national SELINCMedia project. His current research interests include the exploration of heterogeneous media repositories and the role of trust in interpreting and (re-)using media assets.

3. CERTH-ITI Centre for Research and Technology Hellas - Information Technologies Institute
Contact person Vasileios Mezaris (bmezaris@gmail.com)

The Information Technologies Institute (http://www.iti.gr), until recently known as Informatics and Telematics Institute, was founded in 1998 as a non-profit organisation under the auspices of the General Secretarial of Research and Technology of Greece (GSRT), with its head office located in Thessaloniki, Greece. Since 2000 it is a founding member of the Centre of Research and Technology Hellas (CERTH - http://www.certh.gr). CERTH-ITI's related areas of R&D activities include: semantic multimedia analysis, multimedia indexing and retrieval, large-scale and social media analysis, knowledge structures, personalization and knowledge discovery for the Semantic Web, etc. CERTH-ITI has participated in more than 80 EC IST and 85 National projects and subcontracts. Over the last eight years, the ITI research team has authored over 200 publications in scientific journals, 65 books/book chapters and over 500 presentations to international conferences. CERTH-ITI regularly participates in international benchmarking activities, such as TRECVID, and in 2011 had the top performance in the TRECVID interactive known-item video search task.

CERTH-ITI currently coordinates, among other projects, the FP7 ICT IP SocialSensor "Sensing User Generated Input for Improved Media Discovery and Experience", and has leading roles in the FP7 ICT IP LinkedTV "Television Linked to the Web", FP7 ICT IP GLOCAL "Event-based Retrieval of Networked Media", and FP7 ICT STREP PESCaDO "Personalized Environmental Service Configuration and Delivery Orchestration". It also participates in the EDA-funded project MEDUSA and in the audiovisual search engines Coordinated Action "CHORUS+". In the past, CERTH-ITI coordinated or had leading roles in several IST-FP7 or FP6 IPs (WeKnowIt, aceMedia, MESH, X-Media), STREPs (JUMAS, VIDI-Video, BOEMIE) and NoEs (K-Space, Knowledge Web), and also lead the Greek national R&D Training Network MULTI-MINE. CERTH-ITI participates in standardization activities (e.g. MPEG, JPEG), and is active in the organization of scientific conferences, including the Joint ACM Workshop on Modeling and Representing Events (J-MRE'11), the 3rd ACM Workshop on Social Media (WSM'11), the 2nd ACM Workshop on Events in Multimedia (EiMM 2010), the 2nd ACM Workshop on Social Media (WSM 2010), the 2009 ACM Int. Conf. on Image and Video Retrieval (CIVR 2009), and others.

Role in the project

CERTH will contribute to CACHE-22 primarily in the area of multimedia information processing and analysis, most notably on the fragmentation of cinematographic video content, the definition of a light ontology for video analysis and annotation, and the development of the corresponding concept and event detectors for the video. CERTH will lead the related work package WP5 "Video Analysis".

Key personnel

Dr. **Vasileios Mezaris** is a Senior Researcher (Researcher C') with CERTH-ITI (http://www.iti.gr/~bmezaris). His research interests include image and video analysis, content-based and semantic image and video retrieval, event detection in multimedia, medical image analysis. He is the co-author of 22 papers in refereed international journals, 7 book chapters, 2 patents, and more than 70 papers in international conferences. He regularly serves as reviewer of project proposals and of manuscripts submitted to various scientific journals, and organiser, program committee member or reviewer for various international conferences. In 2011, he was one of the General Chairs of the ACM Joint Workshop on Modeling and Representing Events at ACM Multimedia. He has participated in many European and National Projects, having leading roles in several of them. He is a Member of the IEEE.

Dr. **Ioannis (Yiannis) Kompatsiaris** is a Senior Researcher (Researcher B') with CERTH-ITI (http://mklab.iti.gr/mklab_people/~ikom). His research interests include semantic multimedia analysis, indexing and retrieval, social media analysis, knowledge structures, reasoning and personalization for multimedia applications. He is the co-author of 55 papers in refereed journals, 30 book chapters, 7 patents and more than 170 papers in international conferences. He has been the co-organizer of various international conferences and workshops and has served as a regular reviewer for a number of journals and conferences. He is a Senior Member of the IEEE and Member of ACM.

4. IN2 (http://www.in-two.com) contact person Alexandru Stan, as@in-two.com

IN2 is a best-of-class provider of award-winning solutions and services to interface with multimedia content. We transform content repositories into extended media asset management systems by supplying advanced rich media indexing tools and creating flexible interfaces that:

- Enable search beyond metadata and descriptions;
- Allow fast publishing, web 2.0 tagging and expert-level granular annotations; and
- Provide faceted browsing and visualisations of digital collections.
- We specialise in management and publishing of cultural and creative content; organisation of learning resources into educational workflows; multimedia location-based services; and presentation of accessible content. Our core technological expertise includes:
- A service-oriented platform where diverse content-based indexing modules can be composed into pipelines and workflows; easily customisable to store extracted metadata, annotations and other content into any repository structure.
- An environment that exploits repository content by the authoring of flexible user interfaces whereby each interface element can be fully configured regarding its layout, the data element it represents and the functionality it provides. This environment is generic enough to allow for the designing and building of a variety of web front-ends to interface with multimedia content.

Individual IN2 services and service compositions can be licensed to third party media asset management solution providers, purchased as a service on a annual fee basis, or composed together to provide a complete media asset management and publishing system.

Role in the project
Thanks to the expertise in multimedia annotation interfaces and automatic tools for video analysis, the main contribution of IN2 will be in WP6, where together with UNITO and ATC the core system for annotation and search of dramatic media objects will be implemented. In this WP, IN2 will take the lead in designing the system architecture and implemented the integrated software. Apart from this, IN2 will lead the development activities of the interface for human annotation of the corpus (in WP4) and support the video analysis tasks of WP5. Naturally, as an industry partner, IN2 is very interested in bringing the project results to the market and as such will play an important role in the realization of exploitation applications.

Key personnel
Dr**. George Ioannidis**, Director, holds a Master of Science (MSEE) and a doctoral degree (PhD) in Electrical Engineering and Computer Science from the National Technical University of Athens. He additionally holds a Master of Business Administration (MBA) where the main focus of his work was on innovation management. After positions in academia among others as the head of the image and video analysis group at the Center for Computing Technologies in Bremen, he founded IN2 in 2005 and is since then its director. He has initiated, participated and managed several European research projects since 1993 in the area of multimedia content management and accessibility. He acts frequently as a programme committee member in international conferences and workshops (e.g. ACM Multimedia Information Retrieval, International Conference on Pattern Recognition) and serves regularly as a reviewer of business plan competitions, and as an evaluator and reviewer of national, European and international collaborative research projects.

Alexandru Stan is a Programme Manager at IN2 and holds a BSc in Electrical Engineering and Computer Science from Jacobs University Bremen and a MSc in Software Systems Engineering from the RWTH Aachen. He has worked in European projects that have a strong focus on design for all, used a user-centred approach to develop novel mobile content navigational solutions and is managing IN2's participation in them. Furthermore, he is leading IN2's development activities in the areas of multimedia asset management and advanced interfaces to audiovisual content. He has been involved in robotics projects, participating in European competitions. Additionally, he has worked on projects in the areas of databases, embedded operating systems, computer networks and semantic technologies.

Jana Wedekind is a Business Developer at IN2 and she holds a BSc in Media and Communication Informatics from the University of Applied Sciences Reutlingen and an MSc in Digital Media from the University of Applied Sciences Bremerhaven as part of the International Inter-University Study Programme Digital Media Bremen/Bremerhaven. Her academic training focused on Software Engineering, Multimedia Production, Web Design & Development and Media Art. As a web developer she was involved in designing Web User Interfaces, Content Publishing, as well as Online Marketing.

5. Athens Technology Center (http://www.atc.gr), Contact person Nikos Sarris, n.sarris@atc.gr

Athens Technology Center S.A. (ATC) (http://www.atc.gr) founded in 1987, ATC is an Information Technology Company offering solutions and services targeting specific sectors incl. the Media, Banking and Retail Sectors, Utilities and Public Sector Organisations as well as horizontal solutions focusing on Content Management, Web Applications, Human Capital Resource Management and eLearning, and Mobile Applications. The activities of the Company span among several countries in EU, Eastern Europe and CIS countries, as well as the Balkans.

ATC provides the Media Sector with consulting and media – specific solutions as an active and leading supplier of newsroom, editorial and archive specific editorial solutions with a particular emphasis on News and Press Agencies and Publishing organisations. ATC's customers include major newspapers in Greece, Cyprus and Romania and the National News Agencies in Greece, Portugal, Poland, Russia as well as the Western Balkans. The initial engine of the Suite was developed in a Greek State R&D Project (in 1996), whereas new collaboration enabling versions with strong search capabilities are currently being researched in EU Projects. Having acquired ISO 9001 certification since 2000, the company provides a broad spectrum of value-added products and services such as consulting, customer training, installation and maintenance, warranty and post-warranty services, SLA projects, project management, and professional support. Central to the Company's strategy is the conduction of vivid Research and Development, focusing both in improving current Products and Solutions, as well as in exploring new technologies for future growth. All ATC own products and services are based on early prototypes and/or "proof of concept" obtained through R&D Projects, whether funded by the Company, the Greek State or the EC.

Role in the project

ATC's prime role in the project is integration of all system components in a platform that will allow the implementation of the pilot applications. ATC also has commercial interest in the foreseen applications and their business prospects, which are fully in-line with the company business objectives.

Key personnel

Dr. **Nikos Sarris** is a Senior IT Consultant in the ATC Consulting Department. He has received his PhD from the Aristotle University of Thessaloniki and his Master of Engineering degree from the University of Manchester Institute of Science and Technology. He has worked as a Researcher for the Aristotle University of Thessaloniki and the Informatics and Telematics Institute, where he participated in several national and European projects. His research interests include 3D model-based image and video processing, multimedia coding and analysis, and semantic-based knowledge extraction. He is the co-editor of a book in 3D modelling and animation and has authored numerous publications for international journals and conferences. Dr Sarris has been the Technical Director of the MESH IP and the coordinator of the GREDIA, PAPYRUS and SYNC3 STREP projects.

Dr **Ilias Spais** is working as an IT consultant in the Consulting Department. He received the diploma in Electrical and Computer Engineering from the University of Patras in 2000, and the PhD degree in Analysis, Design and Development of Processes, Systems and Computer Engineering from the National Technical University of Athens (NTUA) in 2006. He has been involved in several research projects in the context of the IST framework as a research associate and senior developer of the Telecommunications Laboratory of NTUA. His research interests include natural language processing, speech recognition and synthesis, digital management and presentation of cultural content, distributed information in SOA-based systems and multimedia e-learning platforms. He is currently involved in the field of business applications, user requirements capture, specifications analysis, dissemination and exploitation activities and project management in European and National R&D projects. Dr. Spais is a member of the Technical Chamber of Greece.

6. THE NETHERLANDS INSTITUTE FOR SOUND AND VISION, Contact person Johan Oomen,
joomen@beeldengeluid.nl

The Netherlands Institute for Sound and Vision (NISV) (http://www.beeldengeluid.nl) maintains and provides access to 70 per cent of the Dutch audio-visual heritage, comprising approximately 750,000 hours of television, radio, music and film and web video, making Sound and Vision one of the largest audiovisual archives in Europe. NISV is the business archive of the national broadcasting corporations and is also engaged in large-scale digitization efforts of its analogue holdings. Currently, the digital holdings comprise 3 petabytes; it's expected that this volume will grow to 15 petabytes by 2015.

Sound and Vision has brought thousands of hours of archive footage on-line for various end-user services, including dedicated services for the educational market, footage sales, creative reuse, mobile access. NISV also operates a visitor attraction aimed at the general public, the Media Experience, which is visited by over 200,000 people annually.

NISV is an experienced partner in national and European research projects and active in the international organisations FIAT/IFTA, IASA, EBU and UNESCO. Recent research projects include AXES (FP7), LinkedTV (FP7), LiWA (FP7), CoSyne (FP7), DigiBIC (FP7), PrestoPRIME (FP7), EUscreen (eContentplus) and Europeana V2 (ICT-PSP). NISV has an strong Research & Development department. The department focuses on a number of key topics, including: digital durability ensuring long-term access, understanding user requirements, new annotation approaches, (semi)automatic or using crowdsourcing, contextualization, semantic interoperability and visualisation.

Sound and Vision is host to the EU PrestoCentre competence centre for audiovisual digitisation and digital preservation, and technical co-ordinator of the EUscreen Best Practice Network on access.

Role in the project
Sound and Vision:
- delivers content for training and evaluation
- will evaluate the tools with media scholars
- will disseminate the results of the project within the audiovisual heritage domain, through its involvement in EUscreen, European Television Studies Network, EBU and PrestoCentre

Key personnel

Johan Oomen
Johan Oomen is head of the Netherlands Institute for Sound and Vision R&D Department and researcher at the Web and Media group of the VU University Amsterdam. He is mainly working on externally (FP7, ICT-PSP, CATCH) funded research projects that focus on providing access to digital heritage on the web. Oomen holds an MA in Media Studies. He has worked for the British Universities Film and Video Council and the RTL Nederland.

Erwin Verbruggen works at the Research & Development department of the Netherlands Institute for Sound and Vision. He is in charge of communications for EUscreen and oversees the technical development of the Journal of European Television History and Culture. He also works on a number of projects related to digitisation and search in audiovisual archives.

Roeland Ordelman is senior researcher and senior projectmanager R&D at NSIV and researcher in Speech & Language Technology and Multimedia Retrieval at the University of Twente (PhD 2003). He is co-founder of XMI (Cross Media Interaction), a company that provides services for automatic, speech-based annotation of audio data. He has been working on a number of national and international projects in the area of multimedia retrieval and speech and language technology.

Maarten Brinkerink is project manager at the R&D department. He holds an MA in New Media and Digital Culture and specializes in the distribution of cultural content using digital media. He manages innovative projects like Open Images (an open media platform) Oorlogsmonumenten in Beeld (a location-based iPhone app that enriches war monuments with audiovisual heritage) and Waisda? (a multiple award winning crowdsourcing game for collecting metadata for audiovisual content).

7. Cinecittà Luce (LUCE), Contact person Marco Rendina, m.rendina@cinecittaluce.it

Cinecittà Luce is the most important cinema related public company in Italy, born from the merging of Cinecittà Holding with Istituto Luce, an institution established back in 1924 with the primary goal of using cinema to spread culture and knowledge.

Cinecittà Luce holds today one of the most important European film and photographic historical archives (www.archivioluce.com), in which are collected and digitally catalogued more than 12.000 newsreels, 9.000 documentaries, 3.000.000 photos, and several other titles, ranging from films, dating back to the origins of cinema, to collections and film clips of 20th century events. This is one of the largest audiovisual footage collections in Europe concerning the history of the twentieth century, and it has also been nominated by UNESCO-Italy for the membership registration in the "Memory of the World" registry.

Luce's Film Archive, with a total of more than 4.000 hours of footage, has been fully digitized and it was made accessible on-line since 1998. Today Cinecittà Luce exploits its archive commercially with annual footage sales of more than 2 millions euros.

Within the last 10 years Cinecittà Luce has participated in several EU co-funded projects in the IST, eContentplus and ICT-PSP calls, aiming at developing best practices related to digitization and access in the audiovisual domain. Among them, European Chronicles Online (ECHO), Video Active, European Film Gateway, EUscreen and EFG1914.

Role in the project
Cinecittà Luce will contribute to the project by evaluating and assessing the technology produced in the CACHE-22 project in the context of cultural heritage audiovisual, and film and footage archives, to guarantee the concreteness of the project results and in order to reach the requirements and opportunities of the market in the targeted doamin. Luce will also disseminate project results and research results among film archives, footage libraries and commercial audiovisual archives, evaluating also the impact of the project outputs on the film heritage community at large.

Key personnel

Edoardo Ceccuti. Director of Istituto Luce's Historical Achives since 1991, after three years in the role of Production Manager and Executive Producer of films and documentaries. His experience in the audiovisual sector started in 1973 with Warner Bros Italy; within the company he participated in important national and international productions (such as Amarcord by F. Fellini and Todo Modo by E. Petri); in 1980, in the position of director, he left Warner Bros Italy to join Gaumont Italy where, in 1983, he took charge of the administration management of the whole group. In 1985 he left Gaumont to embark on the activity of independent producer of films and documentaries for the Italian cinema and television market, activity carried on for Istituto Luce starting from 1988.

Marco Rendina. Project manager and IT specialist. Since 2011, he held a researcher position in the Consorzio Roma Ricerche, a university consortium where he was head of the R&D group of the digital libraries and archives division, focusing on all aspects of digitization of historical and audiovisual archives, from metadata standards to database design and preservation policies.

Since 1996, he is collaborating with Cinecittà Luce (and other leading Italian audiovisual institutions, including Cineteca Nazionale), where he followed the development and realization of various digitization and applied research projects. In Cinecittà Luce he has followed several EU co-funded projects like ECHO, Video Active, European Film Gateway and EUscreen, in which he is currently the Network Activities coordinator.

He is actually member of the Italian Organization for Standardization Working Group on Cinematographic Works (UNI GL6) and deputy member for Italy of the European Committee for Standardization - Technical Committee (CEN/TC 372) on Cinematographic Works, where he has been one of the editors of the EN 15907 standard on "Film identification – Enhancing interoperability of metadata – Element sets and structure".

2.3   Consortium as a whole
(No recommended length for Section 2.3 – depends on the size and complexity of the consortium)

Describe how the participants collectively constitute a consortium capable of achieving the project objectives, and how they are suited and are committed to the tasks assigned to them. Show the complementarity between participants. Explain how the composition of the consortium is well-balanced in relation to the objectives of the project.
If appropriate describe the industrial/commercial involvement to ensure exploitation of the results.

i) Sub-contracting: If any part of the work is to be sub-contracted by the participant responsible for it, describe the work involved and explain why a sub-contract approach has been chosen for it.
ii) Other countries: If a one or more of the participants requesting EU funding is based outside of the EU Member states, Associated countries and the list of International Cooperation Partner Countries, explain in terms of the project's objectives why such funding would be essential.

The consortium consists of 7 seven partners, including academic partners, SMEs and content providers in the field of film archives.

The three academic partners cover complementary competences:
- UniTO provides competences in the ontology for drama, the annotation of the drama features, and the connection of the annotation framework to large-scale, common-sense ontologies; UniTO includes researchers in information technologies and drama studies respectively, so forming an interdisciplinary link between the formal representation and the cultural object of the project. Besides directing the management of the project (WP1), UNITO's main contributions to the project will concern: the specification of the ontological framework for the annotation in WP2, where it cooperates with CWI to integrate narrative and media annotation, in the light of the specificities of film archives analyzed by LUCE; the definition of the annotation guidelines and interface in WP4, that will benefit for the interdisciplinary competences of UNITO; on WP6, where UNITO will lead the design of the end user applications for annotation and search/retrieval services, in cooperation with IN2 and the content partners (LUCE and NISV). UNITO will also conduct the dissemination of the project results, especially in the interdisciplinary fields of ICT and humanities. Finally, UNITO will make use of two subcontractors:
    - Since the expertise of UNITO does not cover the technological skills specifically needed to the design, implementation, and maintainance of the web-based tools for the management of the project advancement and the dissemination activities, UNITO will rely on subcontracting.
    - Also, a video repository facility will be acquired from a subcontractor to test the annotation facility, with the configurations required for the testing in WP7.
- CWI provides competences in the ontologies for representing media objects and the representation of the semantic content of a media object, with expertise on use cases acquired through research projects on multimedia resources. Thanks to these competences, CWI will play a leading role in the WP2, bringing its experience on semantic media annotation to the definition of the ontology framework to be employed for developing the annotation platform, and in WP4, where it will be concerned with the use of large scale ontologies on the testbed provided by the corpora selected in WP3 by content partners. CWI will also have a major role in the evaluation of the core features of the annotation facility in WP7. The dissemination activities of CWI will mainly concern the disciplinary areas of multimedia and media studies.
- The IT department of CERTH provides competences in the algorithms for the automatic signal analysis of the audiovisual media objects. In CACHE, CERTH will also contribute to establishing the link between the signal analysis and the annotation of high-level semantic features. In WP5, where CERTH is project leader, CERTH will inform the signal analysis techniques previously devised with the narrative and filmic model incorporated in the annotation framework.

The two industrial partners bring a considerable experience in the semantic treatment of media objects and

ensure the possibilities of exploiting the results of the project. In particular:

- IN2 provides better solutions and services to access the multimedia object via a effective. Thanks to the expertise in multimedia annotation interfaces and automatic tools for video analysis, the main contribution of IN2 will be in WP6, where together with UNITO and ATC the core system for annotation and search of dramatic media objects will be implemented. In this WP, IN2 will take the lead in designing the system architecture and implementing the integrated software. Apart from this, IN2 will lead the development activities of the interface for human annotation of the corpus (in WP4) and support the video analysis tasks of WP5. Naturally, as an industry partner, IN2 is very interested in bringing the project results to the market and as such will play an important role in the realization of exploitation applications.
- ATC provides the content management expertise in archive and editorial solutions, that makes its contribution suitable to the design of the applications for audiovisual annotation and for end-users services. ATC's prime role in the project is integration of all system components in a platform that will allow the implementation of the pilot applications. ATC also has commercial interest in the foreseen applications and their business prospects, which are fully in line with the company business objectives.

Finally, the two content institutions,
- The Netherlands Institute For Sound And Vision (NISV)
- Cinecittà Luce (LUCE)

provide two relevant testbeds for the technologies developed, making sure that the activities are confronted with different materials and different methods of conservation, distribution. Hence the presence of two institutions makes sure that the content we use as testbed on are not tailored to the best practices and standards and laws that characterize national systems. With a tradition in film production and archiving dating back to 1924, LUCE plays an institutional role in the preservation and dissemination of film archives, of high social and historical value. For a decade, LUCE has been specifically committed to the dissemination of filmic archives on the Web, thus providing, besides the content, also the necessary expertise for the definition of the exploitation and evaluation activites on WP7.

Previously involved in EU projects on digitalization and preservation, NISV integrates the role of LUCE with a specific commitment on the design and maintenance of audiovisual archives and their annotation with shareable, standard descriptors in the scenario of open data and cloud computing. For this reason, NIST will play a leading role on audiovisual corpora definition in WP3.

2.4   Resources to be committed
(Recommended length for Section 2.4 – two pages)

Describe how the totality of the necessary resources will be mobilised, including any resources that will complement the EC contribution. Show how the resources will be integrated in a coherent way, and show how the overall financial plan for the project is adequate.

In addition to the costs indicated on form A3 of the proposal, and the effort shown in section 1.3 above, please identify any other major costs (e.g. equipment). Ensure that the figures stated in Part B are consistent with these.

The CACHE-22 consortium consists of 7 partners, committing in total 386 person months over 36 months. This is on average 10.7 person months of effort per month. The total requested funding is M€ 2.5.

CACHE-22 is a STREP, with the goal of increasing access to existing archives of cinematographic cultural heritage. To this end, the project comprises 3 academic partners with complementary expertises in the research areas required for the project goals, 2 SME industrial partners to create the complementary software required and 2 content-owners to supply the content needed to provide example corpora and construct initial end-user services.

Each partner has complementary expertise, and also, given the participation to cooperative research projects in the past, the ability to communicate with other partners to ensure that we achieve our goals.

UNITO has the greatest number of PMs, given their leading role and their contributions to the ontology creation and the core application framework. CWI has the second largest role, given their contributions in both developing the narrative annotations, ensuring the user roles in the design and evaluation of the scenario services, and forming a bridge with the video analysis work. The other partners contribute similar amounts of effort in the video analysis work and end-user application development, creating the software framework and preparing the corpora and the example scenario services.

WPs 2-5 are predominantly research oriented; they advance the state of the art in the semantic models to be employed for the annotation of narrative media, the use scenarios for annotation and search, the signal analysis.
WP3 forms one of the bases of the project, specifying the scenarios that will drive the project, along with the content that will be used as a basis for annotation (WP4) and inclusion in the scenario applications (WP7).
WP2 will combine existing efforts on relevant ontologies and create additions where needed. WP5 results in the (semi-)automatic recognition of filmic concepts in the selected corpora.
WP6 is a predominantly implementation oriented, to ensure the creation of a working demonstration framework that can live beyond the example services developed within the duration of the project.
WP7 creates the services and evaluates them with end-users. These will be 58 PMs of the total 386 PMs of the project. WP7 also contains dissemination activities, in that the user evaluations need to be carried out. These are categorised as RTD, since they are required in the research part of the WP to assess the suitability of the scenario applications for end-users.

Exploitation will be carried out predominantly by the industrial partners IN2 and ATW, and the content owners, NISV and LUCE.

WP1 gathers the management and dissemination activities. Dissemination will be based on the scenario applications, which need to be disseminated to a broad potential user base. Other dissemination will be on the existence of the core application itself. One of the dissemination channels will be the inclusion of the created annotations in the Linked Open Data Cloud. The bulk of the related content, however, will not be made freely available.
The management in WP1 is kept to a minimum. The main effort is provided by UNITO as coordinator, with all other partners having a workload for their management activities corresponding to their overall contribution.

Section 3.   Impact
(Recommended length for the whole of Section 3 – ten pages)

3.1   Expected impacts listed in the work programme
Describe how your project will contribute towards the expected impacts listed in the work programme in relation to the topic or topics in question. Mention the steps that will be needed to bring about these impacts. Explain why this contribution requires a European (rather than a national or local) approach. Indicate how account is taken of other national or international research activities. Mention any assumptions and external factors that may determine whether the impacts will be achieved.

1) HOW THE PROJECT WILL CONTRIBUTE TOWARDS THE EXPECTED IMPACTS
Among the objectives of the Challenge 8, "ICT for learning and access to cultural resources", is the one to "ensure the effective use and exploitation of the cultural resources by developing technologies to make them available, usable and re-usable regardless of their form, location, time sphere etc.". The focus is on:
   • fueling progress in schools, workplaces, museums, libraries and other cultural institutions,
   • investigating technologies that add value and new meaning to cultural digital artifacts and improve user engagement with cultural resources;
   • exploiting Europe's vast and exclusive cultural resources through open and extendable platforms for building services that support use of cultural resources for research and education.
The expected impact of the challenge concerns:
   • Affordability and widespread availability of tools and services for releasing the economic potential of cultural heritage in digital form and for adding value to cultural content in educational, scientific and entertainment contexts;
   • Wider range of users of cultural resources in diverse real and virtual contexts and considerably altered ways to experience culture in more personalised and adaptive interactive settings.

In addition to the objectives for this specific call, CACHE-22 will also contribute to specific goals of theHorizon 2020 vision. In particular, the project will contribute to the integration of humanities into a scientific and technological infrastructure: "Social sciences and humanities research will be fully integrated into each of the general objectives of Horizon 2020." In particular, with respect to the challenges of Horizon 2020, CACHE-22 can have a positive impact on the challenged termed "Inclusive, innovative and secure societies." Understanding and sharing cultural heritage is a key factor to the social inclusion envisaged by the HORIZON program, and innovative access tools to cultural heritage envisioned by CACHE-22 provide for the growth of creative and innovative industries in culture, education and entertainment.
CACHE-22 addresses objectives in the domain of the cinematographic heritage, and of audiovisual narratives in general, by providing an open and extendable web based platform for building services based on the content. As reported in the state of the art, narration is one of the most important vehicles for knowledge and culture, and access to the cinematographic heritage through the key notions of narration (namely characters, actions, events, objects and environments) is a strong cue for the exploitation of the cinematographic cultural resources.
CACHE-22 provides a structure for narrative metadata that can then be used to provide novel means of accessing these resources. The importance of movie clips, such as excerpts of feature films, TV-series episodes, documentaries and newsreels, is witnessed by the abundance of such material in public repositories (such as YouTube). Narratively meaningful clips can be accessed
   • by cultural institutions for cultural dissemination, to promote European cinematographic heritage;
   • by the general public for entertainment, for example finding clips from their own past;
   • by teachers for educational and scientific purposes, for example to explain the narrative structure of a feature film.

CULTURAL DISSEMINATION
Cinema is an important component of the European media industry and an irreplaceable heritage of European culture and history. Cinematographic heritage provides an enormous scale of potential impact. In the core film and TV sector alone:
   • EU FP6 PrestoSpace project found at least 25 million hours of archived film, video and audio material in the 376 archives surveyed in its 2005 Preservation status report. This is the tip of a

growing iceberg, which does not take into account content that was 'born digital'.
- The Screen Digest International 2010 report on "The Global Trade in Audio-visual Archives" found that the trade in content from TV archives was worth M€364 in 2009, and had been growing at nearly 8% p.a. over the previous five years.

Coming to current production, "twelve hundred feature films were produced in Europe in 2010, with almost 1 billion admissions amounting to nearly B€6.5 in box office receipts alone, not taking into account sectors such as home-video, cinema on TV, and Video-On-Demand (VoD) that still represents some two thirds of revenues for film productions. On European TV channels, cinema amounts to 35% of all fiction content; cinema films take 62% of all viewing time spent on pure VoD channels (as opposed to broadcasters' VoD channels). We can thus conclude that cinema is still high in terms of consumption across devices and distribution channels ("When we take into account films on DVD and TV, Britons spent 7551 million hours watching films in 2008. This is 2.4 per cent of available time for work and leisure and 3.7 per cent of available leisure time.). These figures also show how European cinema continues to be a key component in the European media industry." (from the final report of "Digital Agenda for the European Film Heritage - Challenges of the Digital Era for Film Heritage Institutions" – available at http://www.dae-filmheritage.eu/mediapool/100/1000452/data/final_report_en.pdf)

Affordability
In general, Europe is home of the largest collection of film footage, spanning from feature films to documentaries, from the early years of the past century to nowadays. In recent years, and thanks to the digital technology and the Internet, the focus switched from preservation to dissemination. Almost all the relevant institutions in Europe have a public access to their collection (and often are integrated into the European digital library - www.europeana.eu), and, recently, the access to the source is provided directly through Internet. Projects such as the *European Film Gateway* (http://europeanfilmgateway.eu) (a Best Practice Network of 16 European film archives co-funded by the EC under the eContent*plus* programme that aggregate and provide to Europeana more than 700.000 audiovisual items), *Filmarchives online* (the result of the MIDAS project - Moving Image Database for Access and Re-use of European Film Collections - , coordinated by the Deutsches Filminstitut – DIF. MIDAS and initiated in 2006 as a pilot project in the MEDIA Plus programme of the European Commission), or *Europa Film Treasure* (a platform in 5 languages where one can freely watch heritage cinema, both supported by the MEDIA Programme of the European Union), provide a network-platform for on line access, search and retrieval and of the disseminated resource of film and audiovisual resources. Beside the continuous effort to preserve the original audiovisual material, there is a steadily increasing investment in defining a more efficient system to make the documents available.

What CACHE-22 adds to these initiatives is a novel access to the films of the cinematographic heritage: the annotation of segmented clips favors the current trend of postmodern culture to address segments of a work, intended for re-use, re-purpose or in general for fruition, outside of the original context, in a novel sequential order with respect to the original film structure, also hopping from one film to another. Though the navigation of a vast repository of clips is already available in public repositories through tagging, the added value of the CACHE-22 approach is to rely upon the structured knowledge about the narrative structure. This provides both a novel infrastructure for indexing scene clips: clips can be retrieved on the basis of the actions played by the characters, of the characters and objects that populate the scene, of the goals committed by the characters; relying on the narrative structure provides an abstraction over the content description in a cross-media independence, thus connecting screenplays, storyboards, films.

1.	Istituto Luce - Cinecittà: The Archive of the Istituto Luce preserves a vast heritage of film and photographic documents and funds (from 1924, the year of his birth) from private audiovisual collections acquired over time from different sources. Luce's holdings are currently composed of 12,000 newsreels, 9.000 documentaries and other types of films ranging from the origins of cinema to the documentation of events and social life in recent decades.	21.000
2.	British Film Institute (BFI), founded in 1933, is one of the largest film archives in the world, including more than 150,000 films and 600,000 television programmes.	750.000
3.	Bulgarian Cinematheque holds the national film archive, which consists of 15.000 titles with more than 40.000 copies or 300.000 film reels	15.000

4.	Bundesarchiv (Germany)	70.000
5.	Cinematek (Belgium). The film collection has about 143,318 copies of 67,213 different titles: fiction films, documentaries, feature films, short films that illustrate the history of cinema from its origins.	67.213
6.	Cinémathèque française, 40 000 films	40.000
7.	*Cineteca del Comune di Bologna* holds more than 18.000 films.	18.000
8.	Cineteca del Friuli: The Cinematheque was born in Friuli after the earthquake of 1976. Contains about 6000 between fiction and documentary	6.000
9.	Cineteca di Milano: a reposotory of over 20,000 titles of world cinema, from its origins until today	20.000
10.	*DEFA Foundation* stocks include the entire cinematic production of the East German DEFA film studios and further oral history documents of contemporary witnesses. This production consisted in 12,000 films.	12.000
11.	*Deutsche Kinemathek - Museum für Film und Fernsehen*has indexed everything related to film history and technology, cinema and television. It has some 12,000 German and foreign silent and sound films .	12.000
12.	*Deutsches Filminstitut – DIF* holds more than 20.000 film copies.	20.000
13.	Estonian Film Archives (Estonian: *Filmiarhiiv*) holds 8000 films.	8.000
14.	EYE Film Institute Netherlands manages an extensive and internationally prominent collection that features 37,000 films.	37.000
15.	Film Archive in Prague. The film collections currently contain approximately 28.000 copies of feature films, 32.000 copies of non-feature films.	60.000
16.	Filmarchiv Austria ("Austrian Film Archive") is an organisation for the discovery, reconstruction and preservation of Austrian film record material: films themselves, literature about film and cinema, or film-related periodicals. With over 100,000 film titles	100.000
17.	Filmoteca Cataluna. It contains more than 170,000 reels, which include a wide range of film productions in Catalan, Spanish and overseas of all ages. The fund also holds about 9,000 videos and DVDs	260.000
18.	Filmoteca Española a repository of over 35.000 titles	35000
19.	*Fondazione Cineteca Italiana* has a stock of some 20,000 films from all over the world, from the origins of cinema until nowadays.	20.000
20.	Greek Film Archive. The Archive holds the largest and most important film collection in Greece: over 7.500 foreign titles (full length feature films, documentaries and short films), over 2.500 greek titles (full length feature films, documentaries and short films).	10.000
21.	Irish Film Archive is part of the Irish Film Institute the body charged with the promotion and preservation of film culture in IrelandIts collection includes 20,000 cans of film, 5,000 tapes,	25.000
22.	LICHTSPIEL - Kinemathek Bern (LS). The film archives consist of more than 6,000 reels: news, trailers, publicity, documentaries, music films and other short films. Among them are rare and unique films, as well as many amateur films.	6.000
23.	Lithuanian Central State Archive (LCVA) holds a total of 7.612 titles.	7.612
24.	*Magyar Nemzeti Filmarchívum* (Hungarian National Film Archive). The Archive has the most extensive film collection in Hungary, approximately 64,000 inventory items.	64.000
25.	National Library of Norway is responsible for collecting, preserving and restoring the Norwegian film heritage in order to make it available for research and documentation. Its collections include more than 21.000 titles of different formats, genres and origins.	21.000
26.	Royal Belgian Filmarchive (Cinémathèque Royale de Belgique), now called CINEMATEK. The collection now counting up to 100.000 copies corresponding to 45.000 titles: full-length fiction films, documentaries, short films, which illustrate the history of film from the beginnings to our times. The collection grows every year with an average of 2.000 copies.	45.000
27.	Scottish Screen Archive was established in 1976. Principally non-fiction, the Archive's collection to date includes some 32,000 items comprising documentary, newsreel, shorts, educational, advertising and promotional films, amateur and	32.000

	professional productions.	
28.	Young Slovenian Cinematheque faced an important task: to form a collection of fundamental works of the history of world cinema. Till now it was able to collect around 1000 titles.	1.000
	TOTAL (films, newsreels, documentaries)	1.782.825

Although this list does not intend to be exhaustive of all the archives in Europe, the number it produces gives a quite clear snapshot of the ground we can reach with our annotation tools.

To the traditional film archive we have to add the more complex situation of video archive spanning from well established institution (as BBC) to more specialized database (as AHDS on performing arts http://www.ahds.ac.uk/performingarts/collections/index.htm, or TED Talks.). In those case the material cannot be counted in reels on titles, but in hours.

Here are a few of the major institutions:

- BBC archives organized in themed collections of radio and TV programmes, documents and photographs from as far back as the 1930s.
- Britishpathe archive of Documentary, cinemagazine and newsreel that holds 90,000 historic clips.
- Netherlands Institute for Sound and Vision looks after, and provides access to 70% of the Dutch audio-visual heritage (in total, around 800,000 hours of television, radio, music and film)
- The *Institut national de l'audiovisuel* (or INA, French for National Audiovisual Institute), is a repository of all French radio and television audiovisual archives. Since 2006, it has allowed free online consultation on a website called ina.fr with a search tool indexing 100,000 archives of historical programs, for a total of 20,000 hours.
- German Broadcasting Archives The DRA was founded in 1952 as the »German Broadcasting Sound Archive« and renamed the German Broadcasting Archive in 1963 following constant expansion of its responsibilities. In 1994, the former GDR's radio and television broadcasting archive was incorporated into the Berlin location, now Potsdam-Babelsberg.
- SLBA Swedish National Archive of Recorded Sound and Moving Images are currently mass-digitizing parts of its collections, primarily focusing on public service material. Over a three-year period, an initial selection of 1.5 million hours of material will be digitized and made directly available to researchers. Most of the digitization process is automated.
- RAI Teche, the archive of the Italian National Broadcaster RAI, with more than 400.000 hours of television programmes.

The presence of a number of resources and stakeholders in the domain of audiovisual archives makes it possible to envisage the creation and promotion of professional and non-professional, formal and informal communities of annotators, thanks to the availability of the core annotations tool developed by the project. Social web tools, in fact, have fostered a new type of user-generated content, i.e., re-mediated, annotated fan-generated materials, such as transcriptions of films and alternative montages. The project will leverage this recent tradition to promote annotation practices in the general audience.

Wider range
The massive increase in digital audiovisual information is posing high demands on advanced retrieval engines, and it is certain that both the consumer and the professional will need advanced search technology for the management of large-scale video assets. Current search engines, however, mostly rely on keyword-based access, leaving semantic access to the data to research. The CACHE-22 project takes on the challenge of creating a substantially enhanced semantic access to video and film heritage.

Beside the continuous effort to preserve the documents, there is a steadily increasing investment in defining a more efficient system to make the documents available. This represents the growing area on impact of the CACHE-22 project. Internet access, for example, permits a worldwide access to the audiovisual resources, not only limited to film scholars or people with a specific work related interest. With CACHE-22 annotation, the audiovisual documents can be searched and reused for entertainment and/or DIY information assembling. Beside the feature films, there is a heritage of newsreels and non-fiction footage that, if properly annotated through CACHE-22 tools, can give birth to an efficiently accessible audiovisual digital heritage platform, a You-XXth-Century-History-Tube, where anyone can retell the story from a new perspective and according to the subjects and themes preferred.

The list below (information is partially gathered from *filmarchives online*) sums up and gives a brief description of the major institutions and the numbers they hold. A more detailed list can be found at

International Federation of Film Archives.

EDUCATION

Affordability
During the last two decades, Internet has become an important resource for teachers in all degrees of education. The number of sites dedicated to build, collect, and share teaching materials is significantly increasing (cf. Henry Jay Becker, Internet Use by Teachers: Conditions of Professional Use and Teacher-Directed Student Use, Center for Research on Information Technology and Organizations, The University of California, Irvine and The University of Minnesota, February, 1999; Stephen Hartley, Jill Gerhardt-Powals, David Jones, Colin McCormack, M. Dee Medley, Blaine Price, Margaret Reek, and Marguerite K. Summers. 1996.) Enhancing teaching using the Internet: report of the working group on the World Wide Web as an interactive teaching resource. In Proceedings of the 1st conference on Integrating technology into computer science education (ITiCSE '96). ACM, New York, NY, USA, 218-228; Jefferies, Pat; Hussain, Fiaz, Using the Internet as a Teaching Resource in Education + Training, v40 n8 p359-65 1998; Jill H. Ellsworth. 1994. Education on the Internet: A Hands-On Book of Ideas, Resources, Projects, and Advice (1st ed.). Sams, Indianapolis, IN, USA).

Wider range
In trying to add value to existing and, in particular, new records of art, repeating areas of interest are the emotion displayed by the individual actors and the overall configuration of those figures in scenes and sets. Any tools that assist in the automated recognition of these issues will have real impact to both cataloguing and indexing. This is particularly true of the large existing archives where huge collections are being acquired and require cost effective cataloguing which helps identify their artistic merit.
The fundamental value of being able to recognize complex features as emotion is perhaps best confirmed by it being an almost universal element in primary school education. Theatre is one of the principle mechanisms for children to learn about expressing emotion. At its simplest level this is used for developing basic role play in children. Individual works of art are also regularly used for this purpose. There is of course an enormous literature on the role of emotion in art and this too underlines the fundamental value and expected strategic impact of emotional cataloguing tools. Film, TV and even Performance, are more and more interwoven fields that cannot be taught as monodisciplinary. CACHE-22 will contribute to provide a more flexible way of preparing classes for description and analysis of film and drama. Storytelling and dramatic qualities permeate the world of video, movies, gaming and education. The crossmedial approach, together with the capacity of selecting dramatically relevant events in videos and films, creates the opportunity of both an easier job for the teacher and autonomous activities for the students. This will lead to a teach-to-yourself schema in which both teachers and pupils would be more inclined to use the documents.

LEISURE
The above sections describe a specific, although vast, interest of the project in a cultural field. Nevertheless the resource for cultural exploitation can lead to a more leisure and commercial related impact.

Affordability
Broadcasters are increasingly distributing their product on the Internet via specific platform (e.g. BBC I-Player). They offer a web content navigator where audiences can select by channel, by subject, and by title. Given the large repository and often the mixed category (usually some content falls under more than one category), and given overall the different length of content available, it could be fascinating to build our own TV schedule thanks to the augmented annotation of CACHE-22. A similar tool, based on audience comments, for instance, is available on TED website to edit a time custom length of selected "type" of conferences (whether informative, jaw dropping, funny or inspiring). The combination of a superimposed structure of narrative metadata and of content-based automatic procedures that are applied to the video signal, ensures the affordability of the application of the CACHE-22 approach to the vast resources of the cinematographic heritage; also, such a neutral annotation can lead to a number of personalized services, that provide alternative ways of experiencing an already strong cultural resource such as cinema.

Wider range
In principle, CACHE-22 is about unlocking the value of Cinematographic Heritage. The conversion of all sorts of film contents into bits and bytes opens up a completely new dimension of reaching traditional and

new audiences by providing access to cultural heritage resources in ways unimaginable a decade ago. In the future, users of such resources will be able to manipulate digital artefacts online and will be supported by intelligent tools that help them to locate the desired information to create their own works. Therefore, it is essential to develop such intelligent tools that will enable users more conveniently and more naturally to find and combine visual objects they need. CACHE-22 thus aims to contribute towards unlocking this value of visual culture.

2) STEPS NEEDED

CACHE-22 brings together research and industrial partners with highly respected cultural heritage institutions; and in so doing, make the first step needed to reach the impact stated. Indeed, the project will make sure that it will build a strong relation with video and film repositories in Europe through existing European networks. The effective exploitation of CACHE-22 results depends also on a wide agreement among cultural institutions of describing their archives (or parts of them) by means of our tools and on line resources. Such a coordination in plausible thanks to the existing initiatives, such as the Europeana project, which has shown both the possibility of coordination among film and video collection, and a general aim towards a more standard format.

3) EUROPEAN APPROACH

Europe film industry has been put under stress by the continuous predominance of blockbuster USA products that invade each nation with large and complex massive marketing campaign who are culturally undifferentiated. As a consequence this could have a negative effect of moving apart the European audience not only from their actual film production but also from the cultural heritage represented by movies. In other words, the European historic heritage of film and video risks being erased and substituted by a sort of USA colonisation. For example in Europe the myth of the Western quest, is almost as powerful as the one of the fallen empires at the end of First World War. Devising a more functional, accessible and easier way to interpret and discuss our memories is the first step in fighting back the cultural colonization.

Although each nation in Europe has its own national repository of films or/and its archive of video and newsreel, it would be impossible to create a national listing of material without accessing to a large number of across Europe. In fact, each institution held document from its own country but also material from others. A robust European archive cannot be built summing the singled national archive, but only linking and accessing the different organizations. Hence CACHE-22 would not be possible unless on a European base.

4) OTHER RESEARCH ACTIVITIES

CACHE-22 take in account the specific researches on media as LinkedTV project (http://www.linkedtv.eu/), for the TV and Video annotation, and ANSWER project (http://www.answer-project.org) regarding the description of cinematic language. Regarding the analysis of video data we will be linked to the LinkedTV achievements. For the content management we will take into account the outcome of IM3I - Immersive Multimedia Interfaces (http://im3i-server.hku.nl/), NM2 / New media for new millenium (finished in September 2007) The Future of Media Production, and AXES. Overall CACHE-22 will exploit the experience of bringing together institutions and archives as shown in European digital library - www.europeana.eu), and, more recently, web based projects such as *Filmarchives* or *Europa Film Treasure*.

5) EXTERNAL FACTORS

As the research challenges addressed in CACHE-22 are also valid for broadcasters (both national and international), project results can also be transferred to other companies and organizations with similar aims and ambitions. Content providers, such as TV Broadcast Companies, are making their product more and more available for cross platform exploitation. This is reshaping the very idea of film and video fruition. Beside the centralized one-to-many broadcast model, we see the increase diffusion of customized, interactive entertainment and edutainment consumption. If the trend keeps this steady pace, the CACHE-22 model of annotation will play a relevant role for consumer-centred aggregation of content. Interactive TV will include both the idea that the production of new content designed to change based on viewer input (as an interactive fiction), and also, thanks to the CACHE-22 annotation, reuse existing material for, e.g., on-the-fly creation of

alternative plot continuations.

In addition to public broadcasters, archives of audio-visual material can be found in many arenas, including publishing houses, libraries, museums, police and intelligence services, courtrooms that use videoconferencing technology, hospitals, insurance companies, educational services. CACHE-22 technology for indexing, retrieving, browsing and navigating these archives can become valuable tools in these contexts as well.

3.2   Dissemination and/or exploitation of project results, and management of intellectual property
Describe the measures you propose for the dissemination and/or exploitation of project results, and the management of knowledge, of intellectual property, and of other innovation-related activities arising from the project.

DISSEMINATION

The project partners will disseminate the results of this project in a wide range of communities that are active in the heritage field, thus promoting the project results and opening up opportunities for their wider exploitation and for follow-up services and tools.

For the University / Research Centres the project will produce a range of scientific results that will be disseminated through the standard channels and with 2.0/3.0 modes. Results from the project CACHE-22 will be of broad interest to various constituencies, from the ontology research community, who are interested in applications that validate the use of ontologies, to the computational storytelling research community, who are interested in effective representation formalisms for storytelling knowledge and large repositories for extracting such knowledge, to media indexing and retrieval research community, where narrative objects are a thorough tests for automatic approaches that aim at filling the semantic gap.

The following **scientific publications** will be targeted by the consortium:
- **Journals**: ACM Computers In Entertainment, IEEE Multimedia, Artificial Intelligence Journal (Elsevier), IEEE Trans. Circuits Syst. Video Techn. (TCSV), IEEE Transactions on Image Processing, Journal of Web Semantics (Elsevier)
- **Conferences**: Artificial Intelligence and Interactive Digital Entertainment (sponsored by AAAI), ACM Multimedia, ACM Advances in Computer Entertainment, International Conference on Interactive Digital Storytelling (ICIDS, Proceedings in Springer LNCS), ACM International Conference on Digital Interactive Media in Entertainment and Arts; IEEE International Conference on Semantic Computing, AAAI conference, Formal Ontology Meets Industry (FOMI) conference, Formal Ontologies in Information Systems (FOIS) conference, TRECVID, MediaEval.

The project CACHE-22 will also promote a form of **Web 2.0/3.0** dissemination that can involve citizens and heritage institutions in creating a community of annotators of narrative features: the web-based application that stands as the core system of the project can be offered as a basic tool for increasing the size of annotated corpora, and promote the exploitation of the project results. We will publish on the web site a video tutorial that explains how to use the platform; we also invite people to search for items presenting specific narrative features, in order to illustrate the potentiality of the system.

Other channels will be used to disseminate the project results to content providers and the entertainment industry, and obviously to film and video heritage archives interested in having direct access to their content. In the latter case, CACHE-22 will address organizations of the **heritage field** such as Europeana V2.0, EUscreen, PrestoPRIME, FIAT/IFTA, EBU, or the benchmarking initiatives for automatic classification of audiovisuals, such as TRECVID and ImagClef.

Participation in events / Organisation of events / Professional Conferences / Press and Media
The project partners will also address the project results presentations to professional conferences (such as, European Conference on Interactive TV and Video – Euro ITV, and IBC - www.ibc.org) and specialized industrial event of the entertainment industry. Partners will also look for contacts with the local institutions for the promotion of the annotation and search over small local repositories, with typical echoes of generalist media coverage, major newspapers, radio stations and television. In these cases, care will be taken that the message disseminated is consistent with the best practice in scientific research. In local context it will also be possible to organize press conferences or professional meetings with media production companies and media stock companies for the promotion of the annotation of their archives. For example, UNITO can exploit the presence in the local area of the Digital Creativity Hub (http://www.digitalcreativityhub.it/), a consortium of about 60 companies, for organizing presentation events that illustrate the technological pipeline of CACHE-22 and promote the use of the annotation system.

Web site
The CACHE-22 project will promote and maintain a project-specific web site. This site will be populated by background research from various partners in addition to the specific outcome from the CACHE-22 results. In order to reduce the cost and the effort of the site management, the web site will be based on one of the many freely available open-source Content Management Systems (CMSs), such as : Drupal (www.drupal.org), joomla (www.joomla.org) or wordpress (www.wordpress.org).
These systems offer different forms of content distribution (plain pages, picture galleries, downloadable files, blog, forums, polls, ...) as well as a user access control mechanism. The latter, makes these system also suitable to improve the cooperation of the various distributed teams.

EXPLOITATION

There are several initiatives that will be addressed for the exploitation of the project results. We split them into initiatives that are taken by the industrial partners, the content institutions, and the academic partners, resepctively.

Industrial partners

ATC participates in CACHE-22 with an aim to benefit commercially from both the tangible and non-tangible assets that will be obtained. ATC has a long history in developing and marketing content management applications with a series of commercial tools forming the backbone of the company main line of business. NEWS ASSET, DOC ASSET, MUSEUM ASSET and MONUMENT ASSET are some of the products which have been fully developed in ATC and are in the strategic plans of the company to continuously evolve based on new technologies. Exploitation in this sense will take the form of software development contracts by putting in the market enhanced applications that will take advantage of the CACHE-22 tools and technologies. From the perspective of intangible assets, participation in CACHE-22 will grant ATC the capability to offer consultancy services to companies that are interested in asset management applications in sectors where ATC has no particular interest or marketing capabilities. According to the strategy to be adopted during the exploitation activities of the project this may take the form of licensing of software components or contracts for services. The ability to market its expertise in supporting third parties' developments will help leverage the exploitation capabilities of the company.
ATC looks at the market of content management applications as a niche market, where the company can expect to gain substantial operating margins. This is particularly important for all software companies that have experienced significant market erosion in the last few years, due to a significant increase of Offer relative to Demand. ATC will seek to participate actively in dissemination activities, in order to obtain as early as possible during the project, views regarding the wider industry requirements for the technology developments sought. This will enable to tailor an offer that is most suitable for take up soon after the project ends, without excessive post-project productisation effort.

As an SME operating in the market segment of multimedia management and search solutions, also IN2 is very keen on exploiting the research results of CACHE especially the new interfaces for video annotation and search, software for video segmentation and ontologies for media and dramaturgy. The plan after the funding period is to further refine the results, package them and bring them to the market by including them in ONmeedia, their commercially offered platform for managing multimedia content. The strategy for IN2 is both to up-sale to existing customers as well as to acquire new ones by developing new solutions powered by the CACHE-22 results and the know-how gathered throughout the project. CACHE-22 gives IN2 the unique opportunity to strengthen its position in the very competitive market for multimedia management solutions.

Content institutions

For the content institutions, NISV and LUCE, the knowledge on different types of user requirements, best practices and technology that comes from the CACHE-22 project are indispensable components to enable the

exploitation of the rich audiovisual archive to the different user communities, especially broadcast professionals, users of NISV educational platform 'ED-IT!', researchers, and the general public. Furthermore, NISV notably contributes to acclaimed benchmarking tools as TRECVID and ImagClef which arguably set the standards for evaluating technologies relevant to CACHE-22. Being members of Europeana V2.0, EUscreen, PrestoPRIME, FIAT/IFTA, EBU, NISV and LUCE are ideally placed to disseminate the CACHE-22 achievements to a wide community of organizations active in the heritage field. The expected impact possibilities from exploiting CACHE-22 results are vast and all foreseen strategies have been examined thoroughly. CACHE-22 results will help to offer new services for the key user-groups, namely researchers. Sound and Vision has a strategic collaboration with the University of Amsterdam, VU University Amsterdam, Maastricht University and Groningen University. All four include internationally acclaimed media departments. They will be involved in the evaluation of the tools. The universities are member of the European Television Studies Network (http://cms.hum.uu.nl/ethn/) that will play an pivotal role in the wider deployment of the results of the project.

Secondly, by automatically generating a multitude of `access-points' to content fragments that can be used for searching, navigating and linking, the exploitation of NISV and LUCE rich archives will be enhanced significantly. The latter will increase the benefits of the current stakeholders of the archive (broadcast professionals, users of educational platforms, researchers) Moreover, from a Web2.0 point of view, opening up the archive in this manner will eventually automatically create new external access-points to collections as they can be more easily linked to entities in the Linked Data Cloud, stimulating the reuse of NISV and LUCE content.

Besides internal exploitation, we also see external exploitation opportunities. To start with, there are a number of opportunities directly in the context of exploitation within the Cultural and Tourism industries. CACHE-22 automated visualization and annotation tools can potentially reduce the workload for archivists, or at least can be deployed in technology-enhanced manualannotation strategies, allowing archivists to focus on `contextualizing' content rather than on generating low-level descriptions. In the case of NISV, the R&D department will integrate the CACHE-22 technology in the Beta environment they established to test experimental software for internal evaluation.

Knowledge transfer from research institutions

In addition to industrial partner and content institutions, also research partners will contribute to exploitation of the project results, through the education of young researchers and their transfer to industry, being this a valuable knowledge transfer toward some established company or directly by inviting young researcher to start spin-off companies.

The CACHE-22 project supports the creation of spin-off companies thanks to CWI Inc. (http://www.cwi-incubator.nl/), a company established by CWI to provide an environment for researchers to measure their abilities in the business world with minimum early risk, and successful ventures to accelerate on a firm footing into the mainstream of world commerce. The role of CWI Inc. is to be the fertile seedbed where sound new ideas can be prepared to succeed in rapidly advancing, but fiercely competitive, high-tech marketplaces.

Software is also made available to academics and industry for further exploitation after its development in a specific project. Examples of highly successful spinoffs are Software Improvement Group (http://www.sig.eu/en) and Vectorwise (http://www.actian.com/products/vectorwise). SIG has established itself as a thriving company offering software services; Vectorwise grew first as a leading edge technology company before being sold to a large international player.

CWI ensures the distribution of IPR to CWI Inc. and employees taking part in spin-offs, and this is beneficial both to the individuals involved (those taking the personal risks) and to the institute that supported the work being exploited.

The Amsterdam city council also has its own economical development board and encourages the support for small high-tech companies (in ICT and also beyond). (http://kenniskring.nl/)

Recent knowledge transfer in a project with a similar profile was enabled through the transfer of one PhD student with semantic web and user interaction knowledge to a large publishing company. The same work led to the funding of a specific valorisation project where the software produced within the project is now being further developed by a commercial party, so that the original users involved in the research project will benefit from an industrial strength implementation.

On the other hand, UNITO is involved in the Digital Creativity Hub of the Piedmont region in Italy (http://www.digitalcreativityhub.it/). The hub gathers about 60 companies and institutions around the topic of multimedia and digital creativity. Within the hub, UNITO already participates into innovation projects, together with SME's and large companies (such as RAI, the Italian national broadcaster), for bringing research ideas to an engineered implementation, promoting the knowledge transfer to SME's for exploitation. Also, the Turin area includes at least to relevant content institutions, the National Museum of Cinema and RAI Archives, both participating into the Digital Creativity Hub, that are interested into exploiting the results of the project CACHE-22. In particular, the NMC would like to enrich their collections with labels related to locations on the area that have been used as film sets.

Intellectual Property Rights

While CACHE-22 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 CACHE-22 platform will carry an open source license, so that further development by commercial organisations is supported. We will allow the software partners ATC & IN2 to enhance their existing solutions with the main features developed in the project. The exploitation partners LUCE and NISV will also continue to have access to the platforms for CACHE-22 services developed in the project.

The consortium agreement will be the basis in CACHE-22 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
(No recommended length for Section 4 – depends on the number and complexity of the ethical issues involved)

Describe any ethical issues that may arise in their proposal. In particular, you should explain the benefit and burden of their experiments and the effects it may have on the research subject. The following special issues should be taken into account:

Informed consent: When describing issues relating to informed consent, it will be necessary to illustrate an appropriate level of ethical sensitivity, and consider issues of insurance, incidental findings and the consequences of leaving the study.

Data protection issues: Avoid the unnecessary collection and use of personal data. Identify the source of the data, describing whether it is collected as part of the research or is previously collected data being used. Consider issues of informed consent for any data being used. Describe how personal identify of the data is protected.

Use of animals: Where animals are used in research the application of the 3Rs (Replace, Reduce, Refine) must be convincingly addressed. Numbers of animals should be specified. State what happens to the animals after the research experiments.

Human embryonic stem cells: Research proposals that will involve human embryonic stem cells (hESC) will have to address all the following specific points:
- ° the necessity to use hESC in order to achieve the scientific objectives set forth in the proposal.
- ° whether the applicants have taken into account the legislation, regulations, ethical rules and/or codes of conduct in place in the country(ies) where the research using hESC is to take place, including the procedures for obtaining informed consent;
- ° the source of the hESC
- ° the measures taken to protect personal data, including genetic data, and privacy;
- ° the nature of financial inducements, if any.

Identify the countries where research will be undertaken and which ethical committees and regulatory organisations will need to be approached during the life of the project.

Include the Ethical issues table below. If you indicate YES to any issue, please identify the pages in the proposal where this ethical issue is described. If you are sure that none of the issues apply to your proposal, simply tick the YES box in the last row.

<u>Notes:</u>
1. For further information on ethical issues relevant to ICT, see annex 5 of the Guide for applicants.
2. Only in exceptional cases will additional information be sought for clarification, which means that any ethical review will be performed <u>solely on the basis of the information available in your proposal</u>.

ETHICAL ISSUES TABLE

	YES	PAGE
Informed Consent		
Does the proposal involve children?		
Does the proposal involve patients or persons not able to give consent?		
Does the proposal involve adult healthy volunteers?		
Does the proposal involve Human Genetic Material?		
• ° Does the proposal involve Human biological samples?		
• ° Does the proposal involve Human data collection?		
Research on Human embryo/foetus		
Does the proposal involve Human Embryos?		
Does the proposal involve Human Foetal Tissue / Cells?		
Does the proposal involve Human Embryonic Stem Cells?		
Privacy		
Does the proposal involve processing of genetic information or personal data (eg. health, sexual lifestyle, ethnicity, political opinion, religious or philosophical conviction)		
Does the proposal involve tracking the location or observation of people?		
Research on Animals		
Does the proposal involve research on animals?		
Are those animals transgenic small laboratory animals?		
Are those animals transgenic farm animals?		
Are those animals cloned farm animals?		
Are those animals non-human primates?		
Research Involving Developing Countries		
Use of local resources (genetic, animal, plant etc)		
Impact on local community		
Dual Use		
Research having direct military application		
Research having the potential for terrorist abuse		
ICT Implants		
• ° Does the proposal involve clinical trials of ICT implants?		
I CONFIRM THAT NONE OF THE ABOVE ISSUES APPLY TO MY PROPOSAL		