Multimedia Content Management Support in Next Generation Service Platforms

Antonietta Spedalieri, Albert Sinfreu
Telefónica I+D
Via Augusta 177
08021 Barcelona - Spain
+34 93 295 6479
{aspedal, alberts} @tid.es

Pablo Cesar, Ishan Vaishnavi
CWI
Kruislaan 413, 1098 SJ
P.O. Box 94079, 1090 GB Amsterdam
+31 (0)20 592 4332
{P.Cesar, I. Vaishnavi} @cwi.nl

Giuseppe Sisto
Telecom Italia
Via Reiss Romoli 274
10148 Torino - Italy
+39 011 2286161
giuseppe.sisto@telecomitalia.it

Dario Melpignano
NEOS
Via Belfiore 9
20145 Milano - Italy
+39 02 485 91543
dario.melpignano@neosgroup.it

ABSTRACT
The aim of this paper is to illustrate the architecture for multimedia content management support in new generation service platforms. The outlined architecture is based on the requirements identified by a set of interviews with experts from the media industry. According to the results of the interviews, the architecture has been structured into three main areas: content description, content protection, and content delivery and adaptation. The architecture also reflects emerging topics such as end-user content creation and multi-modal user interaction. Finally, it covers the whole added-value chain from the content provider to the end-user. The architecture presented in this paper aims at covering the needs both from content providers and end-users, integrating them in a heterogeneous environment composed by the execution platform by means of different access networks and terminals. From the media industry perspective, the aim of the architecture is to provide a flexible system, so they can offer their content and added-value services. From the end-user viewpoint the aim is to provide an overlay structure that supports a multimedia personalized experience anytime, anywhere.

Keywords
Content management, DRM, Content Delivery and Adaptation.

1. INTRODUCTION
The volume of digital content present on the networks and available to end users is growing every day. The growth of content available is a direct consequence of the increase of the demand: every day millions of people around the world rely on their mobile phone, pocket PCs and computers to access content. They are increasingly using their phones to keep them in touch with audiovisual information, mainly for entertainment. On the other side, the massive availability of camera-equipped mobile devices and other tools for content production transforms end users into content producers capable of creating content that can be distributed and published on the network. These players, professional content providers and end users as content producers, have different requirements on the technical environment that will allow them to produce, publish and search for content in an easy way. Requirements can be summarized as: a standard interface for content description and publication, DRM support for professional protected content, watermarking or other protection schemes to control the distribution of end-user generated content, new charging methods for content access; adding to these requirements the ease of use, the transparence and the non-invasive advertising.

This research has been performed as part of R&D activities in the EU IST SPICE (Service Platform for Innovative Communication Environment). The aim of the SPICE platform is to hide the complexities of the converged communications environment, and to allow commercial services to be developed and deployed in an efficient way [1].

The paper is organized as follows. In section 2 content management objectives and work organization is described in detail. Section 3 presents the requirements co-designed with media experts and end users. These requirements were collected through interviews with key people from the media industry and workshop including potential end users. Their views were also taken into account while refining the system requirements. Section 4 results summarizing the content management architecture and enablers specification are presented and analyzed. This is followed by key achievements and conclusions in the final section.

2. CONTENT MANAGEMENT IN NEXT GENERATION SERVICE PLATFORMS

2.1 Goals of Next Generation Platforms
The main goals for next generation platforms in terms of content management are the following:

MOBIMEDIA 2007, August 27-29, Nafpaktos, Greece / Copyright © 2007 ICST 978-963-06-2670-5
DOI 10.4108/ICST.MOBIMEDIA2007.1790

- To implement a distributed multimedia database, where data and associated metadata can be located; this goal has a major design implication because it takes into account the implications of a pan-European environment composed of various domains;
- To develop an interface that a content provider, service provider, content aggregator and an end user as non-professional content provider would use to manage content;
- To provide the appropriate tools for protecting content by taking into account existing content protection schemes such as OMA [2] specification and additional watermarking methods; This way, the architecture has the advantage of overcoming the intrinsic acceptance hurdles of encryption-based DRM for multimedia content;
- To perform content adaptation and delivery as result of policies based on resource availability and knowledge information such as user context and user profile;
- To give more relevance to the end-user as non-professional content provider, by offering him a kit of tools that let him manage, properly deliver and control the distribution of his content and the possibility of an advanced interaction.

2.2 Innovations

The main innovations presented in this paper are related to the following topics: content selection, content protection, and content delivery. The following subsections address these topics.

2.2.1 Content Selection

Content selection (storage, classification, and searching) focuses on:

- Providing ways to integrate heterogeneous content in a uniform way. This "universal" content guide visualizes stored content ("downloaded or streamed").
- Extending traditional content guide approaches, based on existing content guide schemas, by specifying new ones based on knowledge representation. These new schemas employ the "top-down" approach of the "Semantic Web".

Another innovative aspect worth mentioning is the possibility for the Service Provider to detect relations amongst heterogeneous content available in his Digital Asset repository, in order to better exploit them, and to enrich the content proposition to the end-user. As an example, when the end-user is watching a music video broadcasted on his mobile, the recommendations proposed by the content guide could include interactive links to music of the same artist. Up to now, the end-user has been considered as a passive entity (e.g. watching TV). But, current developments shown by the high increase of P2P traffic and the popularity of services such as Flickr [3], Youtube [4] and Wikipedia [5] clearly show that there is a role-shift in progress; with the end-user becoming both a content producer and distributor. The architecture presented in this paper takes into account mechanisms that allow the end-user to gain control over the content she consumes and to publish self-produced content [6].

In addition to the actual content, another innovation provided by the outlined architecture is to provide the end-user the possibility of generating and publishing metadata. This metadata might be in support of his/her own produced content or to make more accessible for his/her community content produced by other parties (e.g., other users of professional content producers). This article presents new mechanisms for the "tagging" of the content following novel concepts, such as the "folksonomic approach" [7] and the semantic classification. This new functionality is enabled by the provision of a return channel; in particular, this paves the way to interactivity and enables personalization (a metadata database can be queried by the end-user by applying user-defined preferences).

In conclusion, the architecture provides an innovative content guide that takes into account the user as a consumer (profiles, context awareness, and semantic classification) and as a producer/distributor (folksonomies, end-user content).

In addition, the content guide includes both broadcasted and downloadable content. Particular attention has been dedicated to the ongoing activities in the DVB Forum, where they are starting a new topic on "ESG over bidirectional" [8].

2.2.2 Content Protection

The innovation in the domain of intellectual property protection and digital rights management research concentrates on three capabilities:

- Smoothly handling the heterogeneous and multi-device context of the user and of groups of users: the basic approach followed by DRM involves binding licenses, related access rules and content encryption keys to specific DRM Agents following the approach standardized by OMA. A "Domain" approach will allow a license issuer to bind licenses as rights objects, access rules and content encryption keys to a group of DRM Agents, instead of just a single DRM Agent: users can then share protected content, both online and offline, between all DRM Agents belonging to the same domain.
- Supporting IPR in new and emerging content providing contexts, notably with the end user as content provider, via watermarking and ownership assertion techniques that can travel through the different domains without being tied to a specific network, set of devices or content format.
- Supporting offline or unconnected devices content access via storage device supported DRM systems. The system will incorporate OMA encryption into flash memory cards used in mobile phones, to let service providers and mobile network operators offer content based on subscriptions, as well as single-purchase, pay-per-time, peer-to-peer content sharing and gifting, free preview and pay-per-use models.

2.2.3 Content Delivery and Adaptation

The contributions in this area include session mobility, non-monolithic rendering, user-input fusion and end-user content enrichment.

Session mobility provides the means to deliver multimedia content to the user taking into account his context, location, and personal situation. This is provided by the creation of the appropriate decision points that can handle dynamic changes depending on the user situation. Session mobility can be termed as "the content follows the person, instead of the person being forced to follow the content". Due to the number of devices one is surrounded by today and the current necessity of our society to stay connected, it is important to provide the user with a system that supports session mobility between devices.

So far, multimedia rendering has been considered as a one-device process (monolithic). However, now, each rendering device can be considered as one entity, while the actual multimedia rendering can be done in a non-monolithic manner. The architecture presented in this paper uses the open-source implementation of the ambulant player that provides non-monolithic rendering of multimedia content [9].

User input fusion focuses on end-user interaction and should not be limited to the keypad provided by a handheld device. There are other ways of interaction such as voice, haptic interfaces based on position sensors and accelerometers, and pen-based interfaces. Input fusion refers to the process of creating one semantic input derived from a number of instances of the input. These inputs can be in the form of gestures, speech, and traditional user interaction.

3. EVALUATING REQUIREMENTS

3.1 Media expert interviews

In order to evaluate the innovations and requirements presented in the previous section, we interviewed ten media experts in two different European countries. This approach provides an external vision that identifies issues and needs, drivers and hurdles. In addition, we collected general remarks about the value of the network and the requirements among the different actors; with a general scope from the platform technology, R&D perspective to the foreseen implementation and market phases.

Each interview started with a presentation of the motivation and objectives of a new platform for content management and distribution. A specific questionnaire was developed for the interviews. It included aspects such as media expert's vision on future successful services, personalization and social network impact on the actual content management systems and content protection schemes for content distribution. Figure 1 summarizes the topics addressed in the interviews and points out the interest shown by the interviewed in the questions and the relevance of the question to their business.

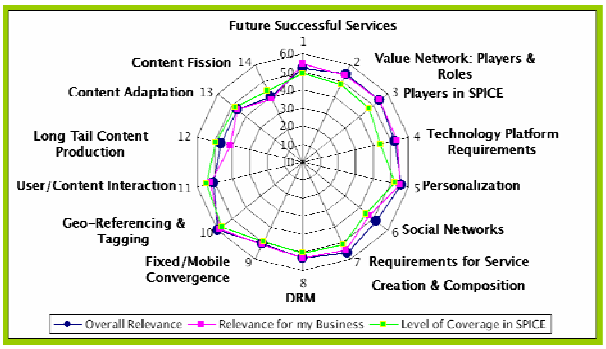

Figure 1: Topics addressed during the interviews with Media Experts and their relevance.

The overall feedback can be summarized as follows:

- More emphasis should be given to the complexity and expected heterogeneity of the emerging business value networks, taking into account more defined roles for the different players and their needs, i.e. in terms of billing and revenue sharing schemes;
- Successful future applications should satisfy practical needs considering that convenience and low cost are key drivers and that services with an audiovisual component are expected to be the most popular service category on mobile broadband;
- Context, geo-referencing and geo-tagging are seen as the most relevant opportunities to deploy content-related services on mobile devices; in many situations context around a piece of content is expected to be of bigger value than the content itself;
- Session mobility, content splitting and user interaction: are seen as interesting functions but difficult to achieve within the currently available technologies; especially splitting of content is perceived as a challenge.
- Content adaptation is not seen positively by Professional Content Providers as they see the content format as a structurally linked component of their product.

3.2 Multimedia Content and End user Interaction

To make the mobile data environment as flourishing as the Internet, we think that an innovative content management platform is needed to provide the end users with a medium to voice their huge communication needs. An added value of a mobile environment is the context information that can enrich content production and fruition. Once user generated content has been produced, it should be stored in an accessible (anytime and anywhere), trusted and secure place. Other content should be made easy to navigate and search. A goal of the platform is to provide a "personal space", enabled with a set of functionalities, needed by the user to manage intelligently his content: storage, protection, sharing, interaction and discovery of end users' content are key functions to be supported.

Incorporating user-generated content is changing the definition of quality database content rapidly, even as search-oriented technologies provide greatly improved indexing tools and navigation structures that often eliminate reliance on traditional database technologies. All these aspects are reflected in the platform architecture for content management described in this paper.

3.3 Requirements Refinement

The interviews lead to a requirement refinement process that ended with the following conclusions:

- As content adaptation is not seen as positive from content producers, the media description will contain the different instances of the content that can be selected depending on the characteristics of the device that is being used at a certain moment; this is achieved by using the multimedia description language SMIL [10].
- End users should be prompted with session mobility or content splitting possibilities: these processes should not be automatic but always happen as a consequence of end user's will. This will be reflected in the implementation;
- Recommendations should be fully controlled by the end users, i.e., users should be prompted with various suggestions, at suitable times only.

4. ARCHITECTURE: CONTENT MANAGEMENT ENABLERS

The architecture for content management defined in the project is shown in Figure 2.

4.1 An architecture to access, store and manage the content

Content guide objective is based on a synthesis of advances in database management system, processing and analysis of multimedia data and artificial and computational intelligence. A great cooperation in these fields has been observed during the last decade. The outcome of this project in this area is a contribution to promote a more effective and efficient access to multimedia information.

In its actual state, the architecture of the Content Guide is able to issue queries at high semantic level to simplify the user access to multiple heterogeneous audiovisual archives, based on different formats and technologies. Another important and considered feature is the management of user's profile metadata that specifies the user's preferences in multimedia content. The content guide architecture is responsible for managing all this user-related data, and uses it for filtering the information returned in response of users' queries.

The Content Guide Service provides information on available contents and services. It permits the user to find information about contents that are available and to select the ones he is interested in.

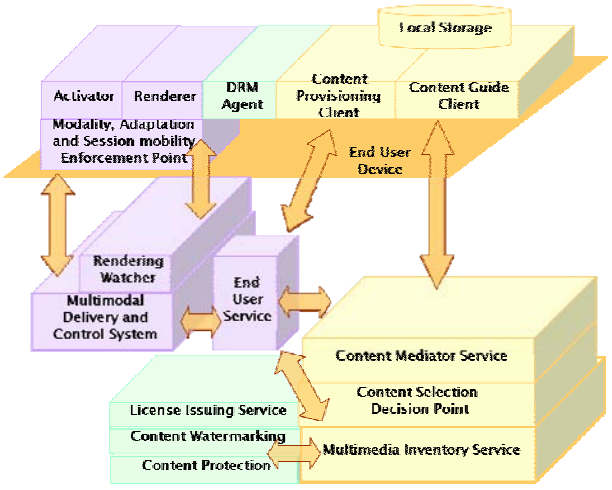

Figure 2: Content Management Architecture

The **Multimedia Inventory Service** consists of a heterogeneous multimedia repository where metadata as well as the content data can be located in distributed data base servers over the network. This enabler is in charge of providing content management functions to other enablers of the platform. It's connected with the **Content Selection Decision Point** that is a wrapper of enablers Query Agent and Metadata Aggregator. They are in charge of searching, managing, selecting and aggregating the various inputs that contribute to the composition of the Content Guide.

The **Query Agent** is in charge of querying metadata repository for content search, matching query parameters against user profile and interrogating the Recommender for recommendations on contents. The **Metadata Aggregator** is in charge of merging metadata, formatting them and generating the complete Content Guide, that is forwarded by the **Content Mediator Service** which is in charge of processing and fulfilling user requests coming from the client application, providing the content guide metadata to the client application, managing the selection of contents, informing the Learning system of user content selections, and retrieving licenses for the selected content. The **Content Guide Client** is an application accessible on the mobile device. It provides the client interface to the user, allowing content search, selection and retrieval. The user interface of the Content Guide can be observed in Fig 3.

4.2 Content Protection Framework

The specification work has been performed with the aim of maintaining independence from specific media objects, formats, operating systems and runtime environments.

In principle, in order for the end user to render the content on her device, she must have both the protected content and the associated rights object, generated by the license issuer. On his side, the license issuer is the entity responsible of granting the permission to render the protected content within the end user device. A critical aspect that has been considered is that the OMA specification refers to the possibility of playing a protected content with a specific device, while the key fruition idea in the platform is based on considering the set different mobile devices at his disposal as a whole. To overcome this limitation a possibility within the OMA specification has been considered and utilized, which is the "Domain" approach, interpreted in the platform as the set of devices within the end user's sphere, which are able to share Domain Rights Objects; these allow the rendering of protected content to the sets of devices that implement a DRM Agent and that can be globally referred as pertaining to a specific end user sphere.

The separation of the rights objects from protected content that enables super distribution and subscription models is therefore the basis for the integration of the DRM requirements within the platform.

The **License Issuing Service** is used to manage licenses that allow access to the encrypted copy of the protected content being accessed together with the **DRM Agent** which, given a protected content URI, plays the protected content. The **Content Watermarking Service is used to** securely embed watermark in content while the **Content Protection Service is used to** encrypt content.

4.3 Content Adaptation and Delivery Building Blocks

The objective of the SPICE project is "to support technically the users by setting-up mechanisms and solutions that provide them with 'always on' and the 'always best-configured' communication environment". The focus here is to deliver and adapt multimedia content based on the current communication environment. This architecture is broken into two different phases:

Figure 3: Content Rendering and content guide prototype.

- Establishment of the communication path between the multimedia content and the end-user. The decisions over the communication path is taken by the resource coordinator based on policies (associated to the end-user, operator, service, etc.), the current state of the communication environment, and recommendations provided by the Knowledge Management Framework implemented in another group in the SPICE project.
- Content delivery: the actual user interaction activity, in which the end-user receives multimedia content and interacts with it.

The **End-User Service** is composed of two inter-related components: **Renderers** (media player) and **Activators** (input handler). The Renderer is in charge of rendering the output (in our case, multimedia presentations), while the Activator is in charge of listening to user interaction (e.g., play/pause/go to next multimedia element). The user input might be a speech command, a gesture, or a more traditional desktop input (e.g., a button click).

The **Fission Component** is responsible, when needed, for splitting a multimedia presentation into composing media elements. These elements can be targeted to different rendering components, through the **Output Transformer** which is responsible for applying content transformations to comply with a given format (e.g., change format or resolution of a video) or even to a given modality (e.g., text to speech transformation).

The **Fusion Component** combines the inputs from multiple **Input recognizers,** which processes raw service inputs generated by an end-user (e.g., in the form of voice commands or gestures). The output of a fusion component consists of service-understandable user input, which is based on the service's input model.

The **Resource Coordinator** is a decision point that chooses on how to use the available resources in a user's sphere (e.g., devices and networks) to deliver a particular service to that user. In addition, it cooperates with the **Session Mobility Enforcement Point** to support session mobility. "Session Mobility" requires that the resource coordinator keeps track of the current state of the multimedia presentation. **The Rendering Watcher** is used to monitor the state of an ongoing multimedia presentation**.** It enforces the timing details, as executed by the Resource Coordinator, and reports the status to the Resource Coordinator.

5. RESULTS: MOCK-UPS

This section includes the preliminary implementation views and explanation of our architecture. It is aimed at explaining how the components defined in this paper work together to provide a determinate functionality: the intention is to visually demonstrate the capabilities and potentiality of the architecture presented in section 4. For this reason, a set of mock-ups have been developed. These mock-ups cover the most relevant functionalities that content management architecture will provide. They also show how the architecture components and interfaces behave for a given storyboard.

We will follow a storyboard to show how the system works, completing it with graphical components interactions. Due to space constraints, only few selected scenes will be showed here.

In the storyboard appears a family with the parents, Philippe and Katherine and their children Patrick and Marie. They live in Paris. They are visiting Barcelona this weekend. All of them are SPICE end-users and, therefore, have a SPICE account. Each has a specific profile created in the SPICE platform with their preferences and user data. We will first demonstrate here how the "Content Guide Client", explained above in this paper, works. This is one amongst many of the SPICE services they available to them.

The family has arrived in Barcelona and is currently visiting the shopping center. Patrick and Marie are tired and not too keen on shopping and prefer watching a movie at the Internet shop, instead. Philippe, thus, decides to download a movie for his children.

Philippe, first, accesses the Content Guide to receive information on the movies available and currently recommended by the movie service provider, according to his and his family's profile and preferences (**Metadata Aggregation** and **Content Push**). Figure 4 shows the interactions between content management components that result in the recommended items showed by the Content Guide client to the end user. In the SPICE platform, a selection of the Content guide client (Step 1), results in a request to the Content Mediator Service (Step 2). The Content Mediator Service is responsible for formulating a relevant query to the Query Agent (Step 3). The query agent then sends a recommendation request to the Recommender system (Step 4) which searches the Multimedia Inventory Service (the multimedia database) for content (Step 6). The received recommendations are forwarded back to the query agent as shown in Fig 4 (Step 6 and 7). The query Agent then forwards this content along with associated metadata to the Metadata Aggregator (Step 8), which in turn forwards this back to the Content Mediator Service (Step 9). The Content Guide client on receiving this content from the Content Mediator Service (Step 10) renders it on the user's device (Step 11) in the form of a Content Guide List.

However, the children want to watch a movie that does not appear in this list, so Philippe performs a query based on the movie title to check for its availability (**Content Searching** and **Profile Matching**). The query also returns prices associated to the available licensing options, such as a license that permits user to visualize the movie on different devices at his disposal for a defined period of time. In Figure 5 the interaction between blocks involved in searching process is depicted. The search that Philippe initiates is passed to the Content Mediator Service, via the Content Guide Client (Step 1, 2). The Content Mediator Service using the Query Agent gathers a set of recommendations on the content using the Recommender System (Step 6 through 9). These recommendations along with

the Content Metadata (incl. licenses) are forwarded to the Metadata Aggregator Service (Step 9). The Content Mediator Service, on receiving the content and associated metadata from the Metadata Aggregator, forwards the Content Guide to the CGC, which then, renders it on the user Device (Step 10).

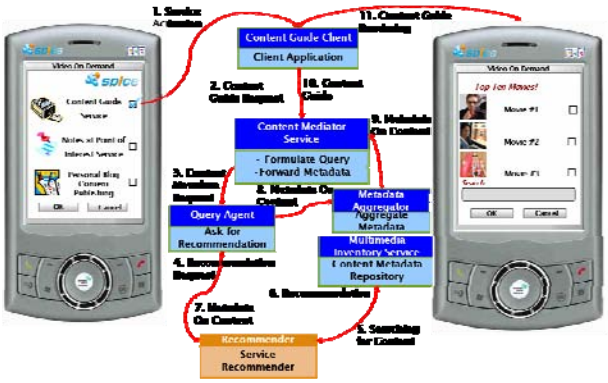

Figure 4: Content Guide is opened on the user device

Philippe selects the license that best suits his needs; in this case the one that allows him to view the movie once on a PC screen or on his mobile device, (**Content Selection**) specifying that the PC in the Internet shop would be the first device that will render the movie (**Establishment of the Multimedia session**). The SPICE platform obtains the movie and pays for the usage on behalf of Philippe.

After the commercial transaction is successfully accomplished, the service provider transmits the movie controlled by the SPICE platform, via ADSL, to a PC in the Internet shop. Based on the capabilities of the selected PC the movie will be adapted in the right way and properly rendered taking into account that a content protected by DRM is being visualized. The functionalities involved in this process are: **content adaptation and delivery** and **protected content rendering with DRM support**. Fig. 6 shows the details of execution in the system for this particular case. It clearly shows interactions of the CGC with a number of other sub-systems, such as End-user Service, Modality and Enforcement Point, presented before in Section 4.

While the children are viewing the movie, the parents finish their shopping and return to the Internet shop before the end of the movie and pause it. The SPICE platform capturing this input, offers a list of possible options, such as: resume the movie now, change device or resume later.

After having received the input from the user (change device), the platform discovers the rendering devices compatible with the movie license in the user's DCS and prepares the transfer of session (**Establishment of the Multimedia session**). The platform, then, recommends the mobile device for rendering the video and the personal audio device for the audio.

After they accept the changes, the platform adapts and redirects the multimedia content (**Service State Control**) to the selected user terminals. This gives the children the possibility of downloading or streaming the remaining parts of the movie to their mobile devices. Fig. 7 shows the details of the system interaction in this case. IT clearly shows how the Media presentation is split into composing media elements and directed to each of the Renderers.

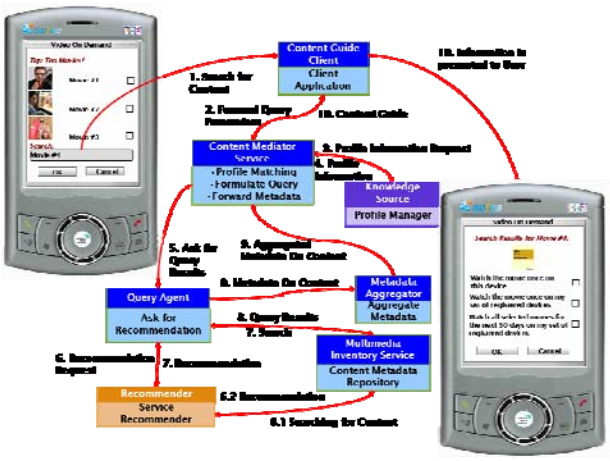

Figure 5: Using the Content Guide to search for content (key-word search type)

We clearly showed, in this section, the behavior of the architecture when the end-user starts the Content Guide Client of the SPICE platform uses it to search for content, selects appropriate content and associated license and finally how the system can transfer media and implement session mobility. The following section presents the conclusions and relevant future work.

6. CONCLUSION

This paper has shown how the content management, protection and delivery functions have to be implemented, to satisfy the requirements that have been formulated and validated, taking into account concrete media industry experts' needs, taking into account a deeper understanding on the state-of-the art of the research areas in question as well as inputs coming from potential end users.

The goal is to realize and demonstrate the novel approach for guiding, adapting and delivering content, in order to enable content-based services across different operator domains, exploit the diversity of content access device capabilities and foster the development of innovative content-based services and associated business models.

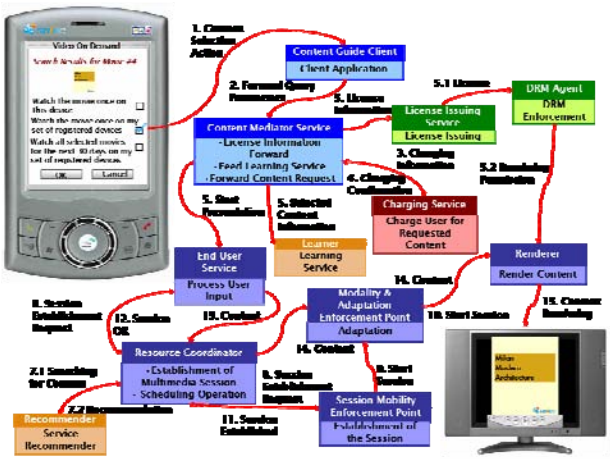

Figure 6: Content Selection, License Issuing, Content Adaptation and Delivery and Protected Content Rendering

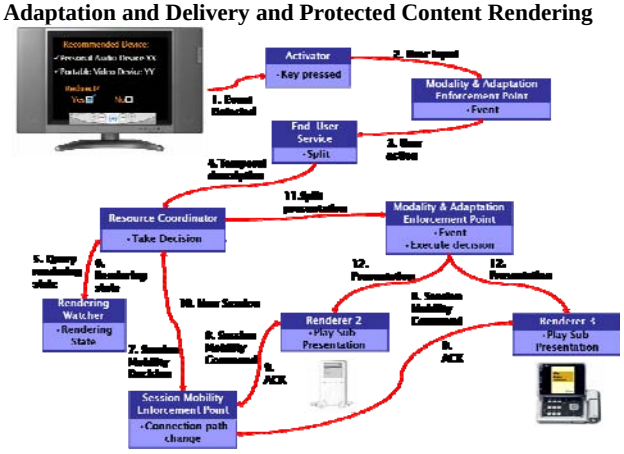

Figure 7: Service State Control, Establishment of the Multimedia Session and Content Fission

The components that form the architecture of the system have been presented and their functionalities have been pointed out and shown in the form of mock-ups and prototypes.

Future work will concentrate on the implementation of content service enablers and their integration with other components, to contribute to the realization of a platform able to support innovative applications and services with respect to content management, protection and delivery.

7. ACKNOWLEDGMENT

The work has been performed as part of the EU IST SPICE (027617) project targeting intelligent extensions for next generation service platforms.

8. REFERENCES

[1] C. Cordier, F. Carrez, H. Van Kranenburg, C. Licciardi, J. Van der Meer, J. Zoric, A.Spedalieri and J.P. Le Rouzic; "Advanced Beyond 3G service delivery environment: the SPICE service platform design principles". WWRF/WG2 Shanghai, 28 April 2006.

[2] http://www.openmobilealliance.org/.

[3] www.flickr.com.

[4] www.youtube.com.

[5] www.wikipedia.com.

[6] P. Cesar, D.C.A. Bulterman, and A.J. Jansen, An Architecture for End-User TV Content Enrichment, *Journal of Virtual Reality and Broadcasting (JVRB)*, Volume: 3, Number: 9, December 2006.

[7] http://en.wikipedia.org/wiki/Folksonomy.

[8] "IPDC over DVB-H: Electronic Service Guide (ESG) Implementation Guidelines"; http://www.dvb-h.org/.

[9] P. Cesar, D.C.A. Bulterman, Z. Obrevonic, J. Ducret, and S. Cruz-Lara, "An Architecture for Non-Intrusive User Interfaces for Interactive Digital Television". Proceedings of *EuroITV 2007*. pp. 11-20.

[10] D.C.A. Bulterman, et al., (eds.), Synchronized Multimedia Integration Language (SMIL 3.0). *W3C First Public Draft*, December 10th 2006