Contents

1 Executive Summary ......................................................................................................................... 3

2 Introduction ....................................................................................................................................... 4
   2.1 UCM cluster Research’s approach ............................................................................................... 4

3 Visions for the Future ....................................................................................................................... 6
   3.1 A review of relevant science-fiction literature ............................................................................. 7

4 Being user-centred ............................................................................................................................ 10
   4.1 Towards meaningful user experiences ......................................................................................... 10
   4.2 Towards automatically personalised adaptive media services .................................................. 12
   4.3 Towards Trust, Privacy, Security & Protection ............................................................................. 13
      4.3.1 Guaranteed service availability .............................................................................................. 13
      4.3.2 Misuse and related liabilities .................................................................................................. 13
      4.3.3 Identification in the Online Environment - Virtual identity card ......................................... 14

5 UCM Challenges towards a Future Media Internet ........................................................................... 17
   5.1 Content aware networks and network aware content/applications .............................................. 19
      5.1.1 Provide an active role to the individuals .................................................................................. 20
      5.1.2 Support of efficient delivery, retrieval and management of 3D data volumes, ....................... 20
      5.1.3 Provide appropriate abstractions content services and how they can form a content network on top of overlay networks ........................................................................... 21
      5.1.4 Support 3D data exchange and interoperation ....................................................................... 21
   5.2 User generated 3D content ............................................................................................................. 23
   5.3 A Future Internet for Collaboration and Interaction ...................................................................... 24
      5.3.1 Interfaces ................................................................................................................................ 25
      5.3.2 Spatial interactive immersive audio ......................................................................................... 26
      5.3.3 Supporting social communication and interaction ..................................................................... 27
   5.4 Immersive media experiences beyond HDTV .............................................................................. 29
   5.5 Users interacting with coupled virtual & physical worlds ............................................................ 30
   5.6 User centric search & retrieval ...................................................................................................... 33

6 Making an Economic & Social Impact in Europe and Worldwide .................................................. 34
   6.1 European IPR ................................................................................................................................. 34
      6.1.1 Trends in managing copyrights ................................................................................................ 34
      6.1.2 Digital rights Management ..................................................................................................... 36
   6.2 An ecosystem for innovation ......................................................................................................... 37
   6.3 Standardisation ............................................................................................................................... 40
7 International collaboration .................................................. 41
8 Conclusions ............................................................................. 42
Annex I - Editors & Contributors ................................................. 43
1 Executive Summary

This document is particularly interesting for readers who want to take part in shaping the Future Media Internet. It considers visions for the future and research challenges on the way towards the Future Media Internet, always with the users in the centre of the considerations.

Internet was initially designed and primarily used by scientists for networking research and for exchanging information between each other. However, due to the explosion of the World Wide Web (which has started as a document repository) and its successful descendants (Semantic Web and Web 2.0), along with the dramatic increase of net-based audiovisual material (networked media) that has been produced by professional and most recently by amateur users, Internet is rapidly transforming to a full fledged virtual environment that facilitates services, interaction and communication. Therefore, the vision that the Future Internet will be an Internet of Media is about to be a reality.

Further, since the combination of media and networking moves in parallel with the user’s drive to acquire more control of their media, the individual “freedom of creation”, understood as a generic ability to act either as a consumer, creator or distributor, is constantly increased. In this context, it is expected that in the near future the User Centric Media concept, which implies that the user will become an active member of the overall media chain by generating, distributing and experiencing high-quality media content, will flourish¹.

Media value chain is evolving towards richer and immersive experiences, where recent advances in 3D processing give rise to innovative applications notably in gaming technologies and in virtual worlds which place new types of traffic demands and constraints on network architectures. 3D collaborative platforms create new requirements in terms of information representation, filtering, aggregation and networking. They also drive demand towards more sophisticated search tools and raise issues of identity management, ownership and trading of virtual digital objects, right of use, and personalised advertisements. These environments coupled with their usage rules hold the promise of a 3D Media Internet which will form the basis of tomorrow’s networked and collaborative platforms². Besides 3D Media, there are other main challenges where the Future Media Internet is expected to evolve: Content aware networks and network aware content/applications, User Generated Content, Collaboration and

¹ “The majority of the 7 billion online videos streamed each month are user generated” (Source Future Exploration Network)

² NEM Newsletter, 17 March 2008
Interaction, Immersive media experiences beyond HDTV, co-existence of virtual with physical worlds and User centric search & retrieval. All of them are described in the sequel and future research directions are given.

## 2 Introduction

This White paper on the User Centric Future Media Internet (UC-FMI) reflects the consolidated opinion of the User Centric Media (UCM) cluster composed of representatives of ongoing FP6 & FP7 EU funded projects, under the guidance of the Networked Media Systems Unit of the Information Society and Media Directorate General of the European Commission. This paper describes the challenges provisioned by the experts for the upcoming years, concerning the FMI, paying specific attention to its provisional 3D characteristics from a user centric perspective. Towards this aim, three parallel, complementary though, Working Groups (WG) were formed:

- **WG1: Personalised & Creative Media**
- **WG2: Future Internet of Immersive Media**
- **WG3: Future Media Internet Communities**

Since the term User Centric Media implies high quality media content generated, distributed and experienced by end-users and taking into account that today's society is becoming increasingly "literate" as a result of technological advances and lower costs in photography, cinematography, 2D/3D graphics design and animation technologies, it becomes clear that the content is expected to play a significant role in the Future Internet (FI). Therefore, the objective of the work of the aforementioned WGs is, by identifying the problems of the Current Internet, to define the main concepts of FMI and to discuss the research challenges that will realize it.

### 2.1 UCM cluster Research's approach

The UCM cluster is currently engaging into a brainstorming session in order to provide a first input to the task at hand and to collect ongoing projects’ experience and the vision of the associated experts, to serve as the starting point and subject for further analysis, elaboration and validation in order to be able to pave the way towards the

---

3 "The views expressed are those of the authors and not necessarily those of the European commission or any of its officials"
FMI and especially the aspects dealing with **Personalised and Creative Media, Future Internet of Immersive Media** and **Future Media Internet Communities**.

With respect to personalized and creative media, the UCM cluster envisions that the FMI will be designed to overcome current limitations and to address emerging trends including: mobility, diffusion of heterogeneous nodes and devices, mass digitisation, new forms of user centred content provisioning, emergence of software as a service and of new models of service and interaction with improved security and privacy features. Moreover, in order to achieve participation and collaboration of professionals and amateurs in a creative way the content repositories and media delivery should not be proprietary, the content should be public, the delivery should be achieved in an unobtrusive way and be access independent.

Mankind has always been trying to reproduce and represent the world around it, using different and increasingly sophisticated techniques and media, from petroglyphs to graffiti, from painting to sculpture, from photography to cinematography, etc. Since their birth, a few decades ago, digital 3D reconstruction techniques have been allowing us to reach an incredible level of realism in the reproduction of the world. Nowadays, the existence of 3D immersive media adds something more to the mere visual reproduction of the world; in fact with 3D immersive media, people cannot only see a representation of the world, but also experiment with and live in it.

Further, given the active role that current, but especially future networked communities will have in transforming society, the UCM cluster finds a strong motivation in conducting research activity in the topic of Future Media Internet Communities. The FI will empower networked communities to radically change the way people communicate, work, learn and enjoy entertainment. Moreover, new media communities will exploit the affordances that the FI will provide for integrating the physical and virtual worlds into new ways of community-based interactions that will empower the user experiences to be more engaging and satisfactory through the provisioning of richer media and more tangible interactions.

The UCM cluster supports that on the basis of what is addressed in the current EU projects and the range of necessary future research activities, the proposed future research challenges in relation to the FMI, could be identified primarily within the aforementioned WGs.

The UCM cluster believes that by supporting European Union goals for prosperity and growth the tackling of all socio-economic challenges is enabled by increasing productivity and readiness for business innovation. This document gives an overview of ICT research areas that are relevant to FMI, for the next five to ten years, from UCM Research’s perspective.
3 Visions for the Future

The development of information and telecommunication services and applications can easily go beyond our normal imagination. Who would have imagined 40 years ago that we are now able to make phone calls, send messages and get our location information with an accuracy of less than a meter with a device not larger than a cigarette box?

Or take Moore’s law that processing power doubles every two years. The effect of these exponentially growing factors is beyond normal imagination. A good example is the price for the Chess game – one rice grain for the first field doubling with every field. Whilst the first half of the board is worth 4,294,967,295 rice grains, the second half is worth about 18,446,744,069,414,500,000 rice grains, more than a billion times as much as the entire first half of the chessboard. Currently we might be just around the middle of the board. Who dares to guess what is about to come …

Some years ago, one of the main science-fiction themes was the teleporting which is referred in literature and cinema. In the Future Media Internet can this be possible? Perhaps the most probable response is YES. Meanwhile applications and services including audio and visual information have been developed; the content should evolve towards covering haptic, tactile and olfactory information for a complete human perception. If we take the 3D visual and 3D audio information with very near accuracy to the resolution and the way the humans perceive these senses, we will have an immersive environment as an evolution of what we have nowadays. If we add the smell, taste and haptic information (which is under research and some partially very good results are already on the market, especially in the area of haptic interfaces), and we can make a good acquisition of the world content (static such as the buildings of a city or alive as humans and animals), achieving the sensatory threshold between reality and fiction. It will be possible to visit any location and talk with the people we want to meet as if they were in front of us? Then, if we reach the necessary content quality, have we not reached the tele-porting ability?

Looking at specific areas beyond normal boundaries, we could end up with issues like:

- **Where can the 3D- and other display technology get us?** Will we finally get the information directly fed into our visual nerve?
- **How will sensory information develop (smell, taste, haptics, etc.)?** This could easily be part of an immersive virtual reality.
- **Person-to-information interfaces.** Will we have direct interaction with the brain through implanted chips? This could also provide means to control external devices, not only to replace/supplement sensory organs and information.
- **How intelligent will the Future Internet become?** Will it become a self-thinking item further-developing itself? Could this finally be dangerous to mankind?
- **What about telepathy as a skill that anybody can learn and practice (may be assisted with artificial implants)?**
To illustrate those “thoughts beyond imagination”, a look at science fiction literature can help.

### 3.1 A review of relevant science-fiction literature

What follows is a first attempt to collect a few views of science fiction literature and movies on how future information and telecommunication services could develop. Maybe this helps starting a process where more such predictions and illustrations can be added.

Concerning the issue where **3D- and other display technology** might get us, we find some stimulating illustrations for example in Star Wars on holographic displays[^4], and in *Time Machine*[^5] on intelligent display panels.

When the first movie of the Star Wars initial trilogy “A New Hope” was released on May 25, 1977, it was a break through in science fiction movies. More than 20 years after the first trilogy, on May 19, 1999, George Lucas released the first of the prequel trilogy, *Star Wars Episode I: “The Phantom Menace”*. Star Wars showed very imaginative display techniques throughout the stories. A very famous scene was the 3D holographic image of Princess Leia displayed by the droid R2-D2. In later episodes various other holographic 3D displays were used. A good imagination how future “video” conferences could look like is the scene where the Jedis sit in a circle discussing various issues. Some of them are sitting in their chairs as holograms participating completely naturally in the discussion as if they were physically present.

Based on the famous novel “The Time Machine by H. G. Wells” of 1895, his great-grandson Simon Wells directed in 2002 a movie where Alexander Hartdegen, a scientist living in 1899 New York City, travels into the future. Alexander stops in 2030 and learns that the Moon is being prepared for colonization. He visits the New York Public Library where he talks with Vox 114, the library's holographic, artificially intelligent librarian. The display is composed of several transparent glass panels on which the librarian walks along with the visitor and answers his questions. Much further in the future, 802,701 AD, parts of the Vox 114 are still working on some broken pieces of the original transparent displays. They brief Alexander on the situation.

**Person-to-information interfaces** are dealt with in various science fiction literatures. An impressive example is given in Dan Simmons “Olympos”[^6], where an old style human “eats” one million books within 45 minutes by reading them directly into his DNA.

“The information began pouring into Harmann now. Information … from a million old books. … This information Harmann was receiv-

[^4]: http://www.starwars.com/
[^5]: The Time Machine, Directed by Simon Wells, 2002
“Olympos” also describes how humans control various functions directly with their thinking. They can for example get access to “Farnet” just by visualizing three green circles within three red squares. Farnet provides information where other people are, a bit like our today’s location services, just with a much wider functionality and an interface integrated in the human body.

The question how intelligent the Future Internet could finally become is dealt with in Dan Simmons “Ilium”.

Several thousand years into the future the situation in the solar system is beyond imagination and control. A post-technological group of humans, pampered by servant machines, begins to question its beginnings. Looking back they find out how intelligent information networks have evolved.

“Thousands of years ago … the old style humans owned a primitive information technology, which they called “Internet”. … Finally it developed to a planet-wide datasphere. But this did not suffice the posthumans. They connected this “Super-Internet-Noosphere” to the Biosphere, the living components on Earth. They connected it with every plant, every animal, and every erg of energy on the planet. Through this they created a complete, total information ecology, which touched everything on, above and under the Earth, a kind of sentient Omnisphere, which only lacked self-awareness and identity.

Then the posthumans foolishly gave it this self-awareness – they did not only develop an artificial intelligence superior to everything, but also allowed it to develop its own personality. … “

This resembles the development of life on Earth. Some hundred million years ago when protozoa started to develop into metazoa, and the individual cells became able to communicate with each other, life on Earth exploded and created an unbelievable variety. The important part was that the individual cells were able to communicate with each other. Through this cells were able to take specific tasks, which is a crucial condition for complicated life forms.

Tele-transporting is an issue happening throughout Dan Simmon’s books “Ilium” and “Olympos”. Old style humans travel by “Faxing”. They travel between “Faxnodes” in no time and without transporting mass. This works by storing all their DNA information in a kind of satellites flying in a ring around the Earth; so only the information is transported and not their physical mass.

---

7 Ilium, Dan Simmons, Eos; First edition (July, 2003), ISBN 978-0380978939
“This fax form of teleportation was staggeringly heavy on computer memory, even with the most advanced Calabi-Yau DNA and bubble memory machines. Do you have any idea how much memory is required to store the data on just one human being’s molecules, much less the holistic wavefront of his or her personality and memories?” …
4 Being user-centred

A User Centric Future Media Internet should make the User pre-eminent in all stages of design and implementation and should lead to a range of new user experiences. In this section we introduce some of the thinking and analysis that designers of the Future Internet will need to adopt in order to efficiently build meaningful new user experiences. We also expose the challenges of delivering personalised media services including trust, privacy, security and protections in a User Centric Media Internet.

4.1 Towards meaningful user experiences

One of the goals of current UCM projects is to provide new user experiences that help in e.g. the nurturing of social relationships. In the field of User Experience research, there are several challenges in developing and designing such user centric media concepts. Whilst effective design is not an exact science a better understanding of how people react to, domesticate and give meaning and value to applications will help designers and technologists to adopt methods and approaches that are more likely to meet with success.

Clever technology design recognises that there is more to product design than simply supporting the user to complete a task. Instead, the design process should focus on developing products that will be domesticated by the users and form an integrated part of their everyday life. Given this context products need to deliver not only satisfaction but also entertainment, enjoyment, a sense of community and positive identity connotations. These qualities are not properties of technology; they are outcomes of the process of a user-product interaction in a specific context and a user’s sense making of this interaction\(^8\). The outcome of this process is often referred to as the User Experience (UX). The shift from user tasks towards ICT as an integrated part of everyday life creates a situation in which the success of a product and the product value for the user cannot be fully understood in terms of usability or Human Computer Interaction theory (e.g. Norman’s seven stages of action\(^9\) and the GOMS model proposed by Card et al\(^10\)).


The user experience is created through a complex interplay of factors such as product interaction, sense-making of this interaction, the predisposition of users (e.g. moods, goals, preferences, earlier experience, etc) and the context (e.g. physical, social, virtual) in which the interaction takes place (Fig.1). The value of technology can therefore only be determined when we broaden our view on innovative products to include not only usability aspects but also these additional processes and aspects as well (as Bannon described so eloquently, we need to start thinking from Human Factors towards Human Actors\textsuperscript{11}).

Especially in developing user centric media concepts, user experience research needs to transcend mere usability research, as media experiences relate to a user’s higher order goals. We can identify three factors that need to be taken into account on each of the aforementioned levels:

- Aesthetic aspects relate to a product's capacity to delight one or more of our sensory modalities and can result in emotions such as thrill, fear, excitement, unease, awkwardness [5].

- Pragmatic & compositional aspects are related to usability, pragmatic and behavioral characteristics of an interactive product and can result in emotions of (mis)understanding of how a product works, (un)predictability of a product and its’ behavior or outcomes and feelings of making progress in a satisfying way.

- Experienced meaning is related to a user’s higher order goals, the ability to recognize metaphors, assign or assess personality or symbolic significance and can result in emotions such as anger, joy, satisfaction, fulfillment, fun, bliss, closeness to one’s own identity or image, inspiration, regret, etc.

The main challenge in the field of user experience research and design is to better evaluate, understand and influence the emotions evoked through product interaction and the interpretation thereof, creating sensible and meaningful experiences\textsuperscript{12,13}.


4.2 Towards automatically personalised adaptive media services

The Future User Centric Media Internet will be expected to deliver personalised adaptive media experiences in real time and with minimal human intervention.

In order for a system to support such innovative services, state-of-the-art mechanisms should be utilised which will be able to map the low-level semantics of user driver requirements to the high-level semantics from the media content. This bridging of the so called “semantic gap” is a huge challenge for the researchers as the extraction of semantics from various type of media content (3D, video, audio, etc) is not a trivial procedure. Each type of content should be processed with specific techniques in order to be semantically characterized.

On the other hand, such services have to be supported by advance personalisation mechanisms in order to initiate the user driven multimedia presentation over multiple network infrastructures. Personalisation enables the multimedia services to be tailored to the needs of specific viewers and to the preferences and perspectives of the media content. A flexible way to achieve the latter is to match the expected and actual live characteristics of the video content to predefined viewer preferences. The personalisation has to be supported on various types of applications. The new user-driven multimedia services will have to develop kinetic type user (task) models in order to enable the end-user, to take advantage of changing their viewing perspective of the event in a way that heightens his/her user experience but yet is simple and intuitive to learn, easy to operate and to remember to use.
4.3 Towards Trust, Privacy, Security & Protection

In the following subsections the horizontal aspects of trust, privacy, security and protection are tailored to meet the needs of the user centric media paradigm. These include guaranteed service availability and privacy, misuse and related liabilities and virtual identification/passport in the virtual environments.

4.3.1 Guaranteed service availability

Guaranteed service availability is one of the key elements of the future User Centric Media platforms. More specifically, the service availability could be divided in two main categories that both of them should be supported:

- Quality of Service (QoS)
- Service Mobility

QoS is the ability of the system to provide different priorities levels on different applications, users and data flows in order to guarantee a certain level of service availability for each of this applications or users. QoS guarantees is important if the network capacity is insufficient and especially for real-time streaming multimedia applications such as online games and IP-TV. All these type of network services often require fixed bit rate and are delay sensitive and especially in networks where the capacity is a limited resource, for example in mobile communications.

On the other hand, Service Mobility is the network mechanism that guarantees the mobility of the services to the end user wherever s/he is located and is transparent to the network s/he is utilizing. All the media services are being provided over the IP protocol, thus the corresponding service for the uninterrupted services provision is the Mobile IP. With Mobile IP a user could have seamless access to personalised services over different networks that belong to the same or different access technology.

4.3.2 Misuse and related liabilities

In many multi-user networked applications (collaborative or not) that do not abide to strict user control and user access protocols (e.g. Massive Multiplayer Online Games - MMOG) any kind of human behaviour may be developed or even thrive, as is the case in real life. In the case of online games and among the vast amount of game players that are advancing their game characters in virtual persistent worlds, some of them might wish to exploit the game services and features in a way not predicted and not desired by the development team and by the game producers and providers.

Among the unlawful or improper game uses are:

- *Exploitation / distribution of design flaws (bug exploitation):* Unintentional or intentional exploitation of game bugs that repeatedly damages another player's character, game play, game economy or game service.
• **Scamming**: any in-game dishonest activity in which one party attempts to cheat another party out of a good, service, or money.

• **Exploitation of Game Mechanics**: Taking advantage of the limitations of the game system.

• **Data mining and data stream manipulation**: Altering the flow of information between game server and client, or extracting information not meant for public (e.g. hacking and claiming other accounts information).

• **Unapproved Third Party software usage**: Usage of any third party software to gain any advantage in-game (e.g. for speeding movement, teleporting the player character or using bots to automatically control the player character).

There is no uniform policy for dealing with these phenomena, and like in real life, the solution is to enforce rules of conduct, the violation of which will lead to a penalty. The policy towards each of the abovementioned unlawful or unwanted game use differs from game to game. Warning, player character deletion, temporary suspension or permanent suspension (player banning) are some of the measures taken to anticipate for these kinds of player behaviour. Extreme actions are also possible but rarely specified in game policies, such as legal prosecution.

Since MMOGs can be seen as a mature implementation of virtual societies, they offer a valuable testbed for the development of optimal policies on misuse control. Many serious synchronous and asynchronous (such as e-Government) applications alike will considerably benefit from innovative mechanisms for inferring trust from a user's actions and interactions.

### 4.3.3 Identification in the Online Environment - Virtual identity card

With the Internet evolving from being “just a game” to an environment where economic transactions take place and “real-life” fortunes are made, it has become more important than ever to be able to create a unique link between users, actions and responsibilities while at the same time preserving the privacy of individuals. Simultaneously, as concerns the protection of minors it is essential for the online provision of services to have effective age control mechanisms. The focus of attention thus, has turned to the means of identification in the online environment, where traditional physical identification documents like ID card, passport or driving license cannot be used. Hence digital equivalents, i.e. “virtual identity cards”, are required.

---

In this regard, there are two fundamentally different options to online identification. Firstly, a user’s virtual identity could be connected to their “real life” identity. This seems to be the approach several European Governments (e.g. Finland in 1999, Belgium in 2003, Austria in 2004) currently follow by equipping national ID documents with smart chips, magnetic stripes, etc that contain information of the card holder (i.e. the identity file which, besides the name and address, may hold for instance the citizen’s registry number or photo) and one or more certificates (which are used to produce a qualified electronic signature). Nonetheless, it is worth thinking of the second option: A user could obtain a virtual identity that is independent of their real life identity and that would be governed by specific virtual laws, i.e., the virtual identity would establish a link to a user’s persona and, by the same token, consequences of their actions would only affect this online persona. Seeing that already today some Internet users strongly identify with their online persona, this idea does not seem to be too farfetched anymore.

Irrespective of the way virtual identification is achieved, however, there is the need for trustworthy authorities that authenticate the user’s identity and set the parameters according to which other entities are entitled to access and process this or just part of this information in order to follow data protection and privacy rules. In this context, the issue of secure communication via the Internet becomes vital. Additionally, the question is whether a virtual ID should reside on a physical token outside of the user’s computer, whether it should be delivered in digital format only or whether it should be a combination of both, the advantage of a physical token being that the user could change computers without too big an effort.

In any case, seeing the different possible approaches, standardisation – not only on a European level – is important. Currently, Article 18 (3) ECT bars the European Union from legislating on matters of passports and identity cards. The issuing of means for identification is left to the Member States, and eventually, there may be 27 different laws in the EU regulating the conditions of identification without any minimum technical and legal standard. Nevertheless, national legislation will at least be subject to

15 For a comprehensive study on all European countries as well as more detailed information concerning the legal implications see the EU document on http://europa.eu.int/idabc/servlets/Doc?id=19132 (last accessed 20th June 2008).

16 The FIDIS (Future of Identity in the Information Society) Network of Excellence is currently conducting an extensive study on virtual identities with the aim to shape the requirements for the future management of identity in the European Information Society, see www.fidis.net (last accessed: 20th June 2008).

17 However, in this context it needs to be mentioned that the European Commission has unveiled a pilot project called STORK (Secure idenTity acrOss boRders linKed) whose aim is to ensure the cross-border recognition of national electronic identity (eID) systems in 13 Member States, so as to enable the cross-border provision of online services. Participating countries are
the provisions under the *Data Protection Directive 95/46/EC* and *Directive on privacy and electronic communication 2002/58/EC* so that any entity that is certified to access the data (primarily the respective public authorities but possibly also private entities) will become subject to the minimum data and data processing quality principles, such as the finality and the proportionality principle (Art. 6 of the Directive). By the same token, the requirements for qualified electronic signature have to be followed, as stipulated in the *Directive on Electronic Signatures 1999/93/EC*.

In summarizing, for the further development of the Internet it is essential that means of online identification to be developed. These means are the only way to help users establish confidence in virtual environments, to fight criminal actions efficiently, to protect minors effectively and to foster eGovernment and e-commerce strategies.

---

5 UCM Challenges towards a Future Media Internet

The UCM cluster identified as main challenges of a FMI the following:

- Content aware networks and network aware content/applications, with emphasis on:
  - Intelligent content delivery; how the media networks will be able to filter the transmitted information based on end-users’ preferences?
  - Content service networks; how to provide appropriate abstractions content services and how they can form a content network on top of overlay networks?
  - 3D data exchange and interoperation; what kind of 3D format is needed for transmitting 3D content over the Internet that can be adapted to various devices?

- User generated 3D content, with emphasis on:
  - 3D content representation with configuration/adaptation capabilities in multiple virtual worlds; will it be 3D video, 3D computer graphics or integration of both?
  - Personalisation and customisation of (3D) intelligent media; how the 3D content will be adapted to the user profile? What will happen if a user belongs to (multiple) community (ies)?
  - Adaptive multimodal content; what kind of technologies is needed for allowing the exact reproduction of the spatial environments? How can we harmonise the transmission and signalling of the multimodal content across networks?

- A Future Internet for collaboration and interaction, with emphasis on:
  - Interfaces; will currently used modes of interaction (i.e. vision, hearing, touch) be extended with smell and integrated with several other sensors and environments (e.g. biosensors, ambient intelligence settings, GPS and other location tracking systems) under a unified multimodal and multi-sensor interaction architecture?
- Spatial interactive immersive audio: is it possible to develop interfaces for immersive multi-channel group communication? Can we achieve a natural interaction between all group members in a high-fidelity surround sound environment?

- Support of social communication and interaction; why and how do people collaborate and interact? How the audiovisual communication can be orchestrated? How the media standards will be extended in media composition, interpretation and delivery for allowing to build social interaction?

- Immersive media experiences beyond HDTV, with emphasis on:
  - Immersivity; how can we overcome the lack of immersivity in the market media applications and work towards a real immersive media? How the existing networked (virtual) environments will introduce an increased social connectivity and an increased immersiveness and seamless media experience of users between multiple (virtual) environments?

- Coupling of virtual with physical worlds, with emphasis on:
  - Online virtual worlds compatible with haptic devices; how the virtual worlds will be accessible using haptic devices? What kind of multimodal interfaces are needed to achieve the latter?
  - Online virtual worlds available everywhere; what types of multilevel architectures are needed for allowing accessibility of the virtual worlds from mobile terminals?
  - Changes in real world reflecting in virtual world; how the virtual worlds will become more secure, allowing for communication with each other and be able to immediately reflect the changes of the real world?

- User centric search and retrieval, with emphasis on:
  - Search and indexing of multimedia data; how will they be performed in a dynamic and scalable way? How the conventional searching mechanisms will be extended and modified to tailor with highly unstructured P2P networks in which new nodes become active/inactive in a continuously dynamic manner?
  - Increasing the content utilization efficiency; how will different tags be derived for different users or different social groups? How the network
will filter out all the unnecessary information incorporating context-aware routing mechanisms?

- Accurate annotation of the information; how the unstructured and non-annotated content will be automatically transformed into metadata structures?

### 5.1 Content aware networks and network aware content/applications

In this section aspect dealing with content delivery, content service networks and 3D data exchange and interoperation will be discussed in the context of putting the user experience at the centre of future research advances.

Content-based routing is one of the key issues for designing an efficient content-aware network. If a "social media routing" could be added to the latter, the networks could be prepared to meet the needs of a User Centric Future Media Internet, where the networks are content-aware, but also User Centric-aware.

**Media networks of the future should be enabled to filter out the transmitted information according to the interests of the end-users.** "Looking inside the data" techniques to allow more efficient congestion control policies, resource management schemes and efficient bandwidth allocation algorithms are necessary, and mechanisms should be implemented to enable it. This way, we achieve a personalised information delivery with respect to the type of the content. On the other hand, content is tagged in a network-centric framework, to allow scalability in content delivery, efficient and personalised searching and indexing, and content filtering with respect to user’s preferences allowing automatic recommendation and media targeting (including advertisement new opportunities).

Networks of the future should be able to:

1. provide an active role to the individuals,

2. support efficient delivery, retrieval and management of 3D data volumes,

3. understand how to provide appropriate abstractions content services and how they can form a content network on top of overlay networks,

4. provide an appropriate file format for transmitting 3D content.
5.1.1 **Provide an active role to the individuals**

In order for the media networks to provide an active role to the individuals, allowing the user acting as both creator and consumer, according to the fact the networks are content aware and the content network aware novel audio visual content handling, routing, managing, coding and processing schemes should be developed for:

a) describing, analysing and managing audio-visual information,

b) efficient 2D/3D media coding in a scalable and adaptable way in terms of network characteristics, terminal devices, context variations and users’ preferences,

c) providing personalisation and context and content recommendation mechanisms,

d) content based handling, navigation and browsing,

e) allowing secure access and guaranteeing copy protection and lightweight rights management schemes in end to end and P2P/P4P networked scenarios,

f) introducing content-based adaptation tools in delivering, managing and manipulating multimedia information,

g) distributing content across a wide range of heterogeneous network platforms as well as terminal devices of different capabilities. These issues should be addressed to establish content-aware functionalities and services to the consumer electronics media devices so that individuals become distributor of their own content across network platforms of different properties and terminal devices of different capabilities to users of different information preferences under a secure and trust environment,

h) upgrading ongoing overlay delivery networks such as P2P and P4P networks (or Proactive network Provider Participation for P2P) with content and context-aware capabilities to optimise the delivery and network performance.

5.1.2 **Support of efficient delivery, retrieval and management of 3D data volumes**

Another issue of research is the fact that User Centric Media networks should support efficient delivery, retrieval and management of 3D data volumes. This transforms the current media 2D world into a virtual 3D world which is more realistic, natural, emotional and representative. 3D data representation enables users to see scenery from multiple views, resulting in a personalised way of watching a media scene, i.e., a simple end user becomes director of what s/he is watching. Optimization mechanisms able to handle 3D data volumes instead of 2D data are required for providing an efficient delivery of the 3D information and achieving a personalised streaming. Personalised rendering mechanisms of complex 3D scenes enable efficient delivery of the 3D
volumes to the end users and minimise the latency experience through the network and associated edge processing platforms.

5.1.3 Provide appropriate abstractions content services and how they can form a content network on top of overlay networks

Apart from the delivery of content, advanced content platforms also have to provide support for various related tasks and processes. These can be provided as value added services allowing the content provider, for instance, to offer a wider variety of formats, easier access or introduce interactivity. Users might be able to access such services for easier navigation, personalisation adaptation of content to their needs, etc. The idea is to use so called content services in conjunction with the underlying network infrastructure to provide a network of content services and by doing so to form a content network.

The goal of this challenge on content aware networks is to understand how to provide appropriate abstractions content services and how they can form a content network on top of overlay networks. Further, how to deal with specific aspects associated with services (such as service discovery, service description, service level agreements) is also part of the research associated with this technical activity. Exemplary services are integrated into the overall infrastructure to fully demonstrate the concept of content service networks. To achieve this goal there are different research challenges to be faced by the research community:

- Content Service Network Architecture and Services Framework, which comprises research on the underlying architectural model for content service networks and also specifies a service framework to easily incorporate all the services making up a content network.

- Quality of Experience: Quality of Service (QoS) refers to the capability of a network to provide better service to selected network traffic over various technologies. QoS is a measure of performance at the packet level from the network perspective. Quality of Experience (QoE) describes the performance of a device, system, service, or application (or any combination thereof) from the users point of view. QoE is a measure of end-to-end performance at the service level from the user perspective. It includes assessment framework allowing building a system for blind (without comparison with reference video), but still accurate video quality assessment.

5.1.4 Support 3D data exchange and interoperation

There is currently a variety of proprietary, closed formats for three-dimensional content. Most of them are created as a file storage solution of 3D data specifically for 3D digital content creation products. Well known examples include the DXF and DWG (coming from AutoCAD, a product of Autodesk Ltd), obj (coming from Advanced Visu-
alizer animation package, a product of Wavefront Technologies), 3ds (coming from 3D-Studio MAX, a product of Autodesk Ltd), lwo and lws (coming from Lightwave 3D, a product of NewTek).

However, there exist already initiatives towards open interchangeable formats for 3D data. Three such efforts that have resulted into finalised specifications and matured implementations are X3D, U3D and COLLADA.

X3D is the ISO standard XML-based file format for representing 3D computer graphics, the successor to the Virtual Reality Modeling Language (VRML). X3D features extensions to VRML (e.g. Humanoid Animation, NURBS, GeoVRML etc.), the ability to encode the scene using an XML syntax as well as the Open Inventor-like syntax of VRML97, and enhanced application programmer interfaces (APIs).

Universal 3D (U3D) file format is an extensible compressed file format for downstream 3D CAD repurposing and visualization that was defined by a special consortium called 3D Industry Forum. The format was later standardized by Ecma International in August 2005 as Ecma-363.

COLLADA defines an open standard XML schema for exchanging digital assets among various graphics software applications that might otherwise store their assets in incompatible formats. COLLADA documents that describe digital assets are XML files, usually identified with a .dae (digital asset exchange) filename extension.

Though most of these formats were made with interoperation in mind, they do not all address the networked applications requirements. All XML specifications do not compress data in the resulted files, while U3D optimises the file size for quick Internet downloading and fast rendering on screen to the expense of omitting most of the engineering data associated with the original drawing.

**Designing an appropriate file format for transmitting 3D content over the Internet** that can be adapted to various devices has many challenges:

1. The format needs to encode not just geometry information but also physics simulation parameters and rendering shaders. COLLADA already does this successfully and is hardware or platform independent.

2. The format needs to be able to encode different profiles suitable for different devices that can connect to the internet. This can be done in current day shader languages that can contain multiple profiles optimized for different graphics cards.

3. The format needs to consider the special needs of a streaming format and must be able to support variable amounts of compression. ASF (Advanced Systems Format) by Microsoft, or Quicktime by Apple are container formats
that are specially designed for streaming but they are designed to handle only video and audio data.

4. Privacy and security are important concerns. Digital Rights Managements is one of the most hotly debated issues in the media industry today and piracy is a global issue. Encryption and watermarking are just two possible solutions to the problem. Also, this will mean that the format is no longer open but that is not desirable either. Intellectual property concerns have to be clearly addressed when any such standard formats are adopted.

5.2 User generated 3D content

Nowadays that content has been established as the most valuable active of the Internet value chain, user generated 3D content is expected to be one of the major research challenges to achieve the objectives of the User Centric Future Media Internet in terms of more attractive and human-friendly content and user immersivity.

Towards this aim, the following main topics can be identified:

a) 3D content representation with configuration/adaptation capabilities in multiple virtual worlds: 3D Content representation should be based on two technologies: 3D video technologies and 3D computer graphics which should be linked together to produce a real 3D complete representation. Virtual reality should integrate both, with standardised and optimised media exchange mechanisms and with configuration and network, terminal and user adaptation mechanisms.

In 3D graphics, the personalisation and adaptation mechanisms are nowadays in the research domain, but 3D video adaptation and personalisation is a challenge that should be strongly supported. Every part of the content distribution chain should be improved: the video coding technologies (e.g. from SVC/MVC/DVC/RVC concepts applied to 3D video coding), or the semantic adaptable metadata representation to the self-configurable intelligent 3D content are some of the examples of the long way to reach a real 3D content representation with configuration and adaptation capabilities.

b) Personalisation and customisation of (3D) intelligent media: personalisation of 3D user centric intelligent media needs the support of context-aware techniques for information filtering which can adapt the 3D information to the user profile (when acting as a consumer in a community). Intelligent 3D content should be based on advanced hybrid methods of collaborative filtering, very suitable when users are members of communities.

c) Adaptive Multimodal content: Traditional media content is evolving towards a more interactive and immersive capabilities and Future Media Internet should not be out of this movement. In order to enhance the multimodality it is needed to develop self-adaptive intelligent user content adaptation technologies with multimodal capabilities.
(such as haptic interfaces) which can interact with remote multimodal interfaces (as for example in the user premises).

There is no standardised (or market accepted) way of providing this multimodal content in nowadays networks and there is an opportunity to harmonise the transmission and signalling of the multimodal content across networks.

FMI is expected to cover haptic interfaces in the content services and applications as it is done nowadays in the console gaming market while being self-adaptive to the heterogeneous terminals where content will be consumed in the Future Internet domain.

3D spatial audiovisual experiences with fully user interaction are the base for novel media services based on user interaction with the environment. In the past, mixed and virtual reality was developed for the user interaction with 3D environments allowing a spatial interaction helped with tactile devices. But Future Media Internet should go beyond these concepts towards a global 3D spatial interactive audiovisual environment irrespective of the user location and which can lead to a real pervasive interaction.

3D technologies can reproduce virtual environments which can be translated and adapted to the user location. With a working Future Internet this environment can be adapted and socialized among peers or communities creating an environment suitable for interaction.

Then, 2 steps should be followed:

a) Developing technologies for allowing the exact reproduction of the spatial environments (as it was described).

b) Improving the technologies from a human-centred perspective for a better 3D representation of the environment. From a human-centered perspective, it becomes clear that the following main challenges can be put forward: 1) the design of meaningful semantics, taxonomies and folksonomías in order to easily create and communicate with 3D objects; 2) the large adoption of such semantics, especially in different cultures; 3) the design of the right level of adaptivity and personalisation of each system; 4) the design of emotionally-charged 3D objects and the methodology of the evaluation with users.

5.3 A Future Internet for Collaboration and Interaction

Whilst telecommunications have made it possible to shrink distances and allow people to communicate from afar, any future Internet should seek to enable much more. The goal should not be just real time communication but the capability to enable shared experiences which are as compelling and as rich in their communication as are real
world interactions. This section looks at a number of key challenges that we believe are critical to enabling such experiences.

They include the challenge of designing user interfaces that are more natural for users, the challenge of spatialised and immersive audio, the challenge of higher fidelity visual images and the challenge of representing and framing experiences to meet users’ social needs.

5.3.1 Interfaces

Most contemporary applications are designed and built in logical tiers that separate the functions of presentation, of business logic and of storage. The top-most tier, presentation, is what the user sees. Application interfaces, as the cornerstone of that tier, affect the usability and user satisfaction of using a product and its commercial success. Providing products featuring a more human interface contributes to their market potential and success. Multi-modal interfaces approach this paradigm. Current state-of-the-art of multi-modal interfaces has moved from dialogue systems to multi-modal interaction frameworks, where the processes for conversational systems based on natural language are augmented with semantic-level content, such as that provided from visual input modalities.

Computer and console games, which are used as an application test bed for the massive commercialisation of many technologies (i.e. 3D hardware acceleration, rendering engines) might be the first application genre to embrace innovative interfaces for interaction and game play control. Many new tactile (joysticks, trackballs, gamepads) and even haptic (force-feedback gamepad directional control and driving wheel control devices) controls have been emerged to provide alternatives to the traditional keyboard and mouse means of input but they have not explored the full potential of existing multimodal interaction technologies. Innovative concepts such as motion tracking, face recognition, multi-directional voice location tracking (as featured in Sony Playstation’s EyeToy, its Eye successor and in Nintendo’s Wii remote wireless controller) coupled with smart applications that make good use of them have led to successful products especially among families and people of all ages and genders. This fact indicates that the use of multimodal interfaces apart from enhancing the user control exerted to any application, can render attractive applications intended to special user groups that otherwise would not use ICTs, such as technology illiterates, children or to the elderly.

Beyond the current state of the art, we foresee that the currently used modes of interaction (i.e. vision, hearing, touch) will be extended with smell and will be integrated with several other sensors and environments (e.g. biosensors, ambient intelligence settings, GPS and other location tracking systems) under a unified multi-modal and multi-sensor interaction architecture. Such an effort for multimodal control of web content has been initiated by W3C Multimodal Interaction Working Group (MMIWG) and the specification for such architecture has been laid out as a working draft (cur-
Currently in its fourth release), but it is not yet mature to become a W3C recommendation, and thus there are not any compliant implementations as of now.

Existing research results and experience gained from the FET European programs in Presence will considerably boost the European progress and expertise in multimodal interfaces.

5.3.2 Spatial interactive immersive audio

It is recognised that the perceived quality of multi-layered representation and composition of rich media experiences is significantly affected by the quality of the audio reproduction for each user. This is relevant to an asynchronous experience, in which a user receives a composition made from pre-existing media components, because the combination of music, effects and speech has a critical creative role in defining the mood and continuity of a narrative. However, it is even more relevant to a synchronised experience constructed between multiple physical environments because social interaction between two or more individuals relies heavily upon them being able to hear each other’s speech clearly and consistently.

The overall audio communication chain can basically be divided into three major components: Recording and rendering of spatial audio (objects), audio coding, and the transport module (FP7 project TA2).

![Audio Communication Engine of the FP7 project TA2](image)

**Figure 2: Audio Communication Engine of the FP7 project TA2**

The major goal of intelligent audio recording is to develop an acoustic interface for immersive multi-channel group communication. The interface should allow for a natural interaction between all group members in a high-fidelity surround sound environment. One focus lies in developing recording techniques that are suited to object-oriented management of multiple participants and audio channels. This approach is in line with the concept of spatial audio object coding (SAOC) which enables efficient transport and interactive rendering of different audio objects. The methods will
explicitly take the characteristics of the human auditory system with respect to spatial perception of sound into account, rather than blindly follow estimation theoretical principles. The acoustic environment is described by means of directional cues and complementary non-directional information. This very general approach to the representation of spatial audio is also beneficial for realistic rendering of the corresponding surround sound on the receiver side using arbitrary loudspeaker setups. The directional information gathered by the audio acquisition unit can also be used to support intelligent user feedback, e.g., in order to localize and to anticipate areas of interest during a gaming session. The microphone setup used for determining the directional information has to be carefully designed in order to ensure robust and accurate results, while remaining as unobtrusive as possible.

5.3.3 Supporting social communication and interaction

Numerous approaches for collaboration and interaction currently exist, and the challenges facing each are different. The traditional technology-led view relates to the massive deployment of collaborative and participative 3D worlds within which people can interact through their Avatars. An alternative view (favoured by the FP7 project TA2), is to use Internet based technologies to support real world social activities. The two approaches are not mutually exclusive, but each investigates solutions to the problem of effective collaboration and interactivity from different perspectives. From a human-centred perspective, a major problem concerns the design choices to be taken into account before presenting some parts of the system as virtual ones, and some others – as real ones. Another major question concerns the interaction means and devices to be used in order to effectively and pleasantly navigate in mixed environments. Yet another issue is the accessibility of current and future interaction devices as they will be used by a very large number of users.

Why and how do people collaborate and interact? In a work context, collaboration and interaction is necessary to achieve goals. In a social context the objective may not be so obvious. It can be argued that collaboration and interaction provide us with opportunities to meet some of our ‘higher order’ fundamental human needs, described by Maslow\(^{18}\) as our needs for love and for esteem. Collaboration and interaction between people allows us to feel a sense of belonging to a group (and so help us meet our love needs), to illustrate our achievements and capabilities (and so give the opportunity to meet esteem needs). It is important to note that these needs, unlike the physiological needs for warmth and shelter can only be met through communication with other people. They are most easily met when that communication is clearly honest, otherwise we are left with ambiguity and may fail to meet our needs for esteem or love. Mehrabian\(^{19}\) observed that in communicating with people, especially about feelings or thoughts, non verbal cues such as body language and tone of voice were important as, where a conflict existed between what was said verbally and non-verbally, people

\(^{18}\)“A Theory of Human Motivation” A.H Maslow Psychological review 50, 370-396 1943

\(^{19}\)“Silent Messages” A Mehrabian, 1981 edition, pp75 -80 and references numbered 286 and 305 therein
most often believed the non verbal communication. In “real life” then, in order to satisfy our needs for love and esteem we naturally seek interaction and collaboration and within these processes we need opportunities to communicate using verbal, vocal and visual cues.

In social situations we invent, or subscribe to, framing events that allow interaction and communication to take place; it might be a social gathering, a meal with friends, a class used to improve a skill, it may be going for a walk, or a cup of coffee. All are plausible social activities that may lead to our higher order needs being met.

In order for any future Internet to take a significant role in such activities, it is clear that the technology should be able to become part of a framing event that is attractive to the participant and it must also support communication at verbal, vocal and visual levels. Unless these criteria are met then we should expect there to be little for people to gain through taking part. With no motivation to take part in this way any fledgling application should fail.

So what are the challenges for any future Internet in delivering such capabilities?

The three core technical challenges are:

- **Orchestration of the audiovisual communication**: where the goal is to improve audio visual communication on any future Internet we need to improve the way the audio visual communications is orchestrated. We need the system to automatically “call the shots” and to mix the result into an attractive and complete representation of the interaction. To do this in real time depends upon the ability to reason with and act upon the interpretation of the multimedia feed.

- **Multimedia interpretation**: The orchestration of the audio visual communications refers to the reasoning processes whereby decisions are reached regarding which audiovisual information will be captured from, and presented to, the different sites involved in the interaction. Multimedia interpretation refers to the extraction of information from the surveyed and/or captured media, to be used as input for the reasoning processes subsumed by orchestration. It will include the extraction of audio and visual semantic cues to enhance and capture the experience, and the detection of events and trends using the audio and visual cues.

- **Multimedia composition and delivery**: Earlier we posited that any attempts to build social interaction within any future Internet would fail unless the framing activities were sufficiently attractive. Part of this attractiveness will be down to the way the event is represented on the screen; to do this effectively will need extension to standards in media composition and delivery.

In summary, we see clear challenges in the support of real time collaboration and interaction, we see these as extensions to established video communications capabilities and expect them to include multimedia composition and coding standards, audio-visual orchestration and multimedia interpretation. In addition technologists need to
be constantly reminded about the nature of human beings so that technology investment can be focused on activities that are likely to meet the needs of more than the inquisitive nature of an engineer's mind.

5.4 Immersive media experiences beyond HDTV

The horizon of immersive media experiences is being explored by several innovative media artists. Indicative of this fact is the Resonance Project activities, where at-location and remote performers are intelligently mixed in a live show. Interesting to this project is the use of a "performance as research" model, within which scientists and artists collaborate to explore a re-visioning of cyber culture and corporeal presence.

Apart from image, sound immersion has been explored in digital production, but has not been ported to the other sectors (e.g. home entertainment) with success, mainly due to its high cost. Precise control of spatialised audio may considerably enhance the usability and usefulness of many 3D applications, including applications such as blind people navigation and location awareness.

Future Internet is the right place to overcome the lack of immersivity in the market media applications and work towards a real immersive media. One of the main supports in the market that immersivity has nowadays and it is expected in the future, are the video games. User immersive experience has been researched from several years ago. The building up of the Internet capacity, jointly with the media advances in 3D media and “HD and beyond” can produce a real immersivity. The proliferation of peripherals (mainly in video gaming) attached to the user terminals (such as PCs) can be used to produce the same results irrespective of the user location and can be adapted to the context and terminal.

Furthermore, enhancement of existing networked (virtual) environments will introduce an increased social connectivity and an increased immersiveness and seamless media experience of users between multiple (virtual) environments. ‘(Virtual) worlds’ will be easier to experience, more entertaining and more immersive than simply viewing images on a PC monitor or (HD)TV.

Based on existing state-of-the-art the central vision is that by making the required technology unobtrusive and distributed, it can effectively ‘disappear’ away from direct user’s attention but still be present in the background (being ambient, i.e. surrounding the user). This allows users to focus more on the content/experience and less on the technology and the ‘technical’ interfaces to control it.

Examples are ‘ambient devices/systems’ such as ‘AmbiLight TV’, ‘turning a home’s glass windows into immersive displays’, ‘amBX’, ‘light speakers’, ‘3D displays’, ‘User Interfaces like the robot iCat and ‘the uWand remote control [for point & shoot type of control]."
5.5 Users interacting with coupled virtual & physical worlds

Tightly coupling virtual and physical worlds is a theme that has captivated the imagination of scientists and artists for a long time, as it is evident by its frequent use in science fiction cinema\textsuperscript{20} and books. All of these are based on the idea that we, as humans, will be intrinsically connected to our virtual worlds.

Since now, art, literature, radio, TV, cinema mainly have been offering representations of reality in which it's possible for the audience to be involved thanks to the identification with the main characters. Our real life is often set apart in order to make space for the life of the characters.

On the contrary, the birth of 3D together with the opportunity to enjoy 3D contents in real time amplifies the concept of user at the centre of media, bringing this concept to its zenith. This can be considered a turning point in the transmission and fruition of contents.

As we step towards an age where the ready availability of information and knowledge is no longer a concern, we seek more and more to connect with this virtual digital world. The dream is to network and inter-connect the virtual and the real worlds so seamlessly that the transition from one to the other is transparent.

Research in cybernetics, networking, artificial intelligence, computer graphics, psychology and neurosciences are all moving us toward interfacing better with the digital world.

The first steps to achieving such a coupling have already been taken. Physicists have coupled a real and virtual pendulum that can now swing together and automatically adjust to move at each others frequency\textsuperscript{21}. On the other extreme, neurophysicists have spliced neurons with semiconductor chips in order to create a brain-machine interface\textsuperscript{22}.

---

\textsuperscript{20} Ghost in the Shell, Mamoru Oshii, 1996.


Real time virtual environments integrated with Internet lay the basis for the perfect coupling of virtual with physical world; this is probably the beginning of Web 3.0 or Web 3D.

“Three-dimensional environments are greatly appreciated by some users and are helpful for some tasks. They have potential for novel social, scientific, and commercial applications, if designers go beyond the goal of mimicking 3D reality.”

Coupling the real and virtual worlds is one of the most attractive and challenging research domains that the user centric future media internet cannot afford to ignore; we can identify several challenges for Future Internet in this specific field:

- **Online virtual worlds compatible with haptic devices**: at the moment virtual worlds are mainly accessible and explorable simply using mouse and keyboard; obviously in order to have a better coupling between physical and virtual, the online virtual worlds should be compatible and experienced by the user by means of haptic devices which enhance the engagement of the users, involving not only her/his head and hands, but also her/his entire body. (As it has been happened with any other technology, the gradual diffusion of haptic devices will probably lead to a significant decrease of their cost, with the consequence of augmenting their diffusion further on).

- **Multimodal interfaces**: The majority of interactive systems still use the keyboard and a pointing device, such as mouse, for input and are restricted to a colour display screen with some sound capabilities for output. System development is usually based on adding more and more visual information to the screen. As system becomes more complex, the visual channel may be overloaded if too much information is presented all at once. This may lead to frustration or errors in use. **Multimodal interfaces process more than one combined user input modes in a coordinated manner with multimedia system output**. Input modes include speech, pen, touch, manual gestures, gaze, head and body movements. These interfaces incorporate one or more recognition-based technologies and recognize naturally occurring forms of human language and behaviour. In some situations people temporarily are unable to use a particular input mode, for example while driving a car, user cannot use manual input but speech is available. Multimodal interfaces provide the adaptability that is needed to accommodate the continuously changing conditions of mobile use. The natural alternation between modes

---


that is permitted by a multimodal interface can be effective in preventing overuse and physical damage.

- **Online virtual worlds available everywhere**: mobility is a hot topic nowadays, the demand for applications to be used also in a nomadic situation is incredibly increasing, and many researches are focused on this topic (for instance Marin Soljačić, an assistant professor in physics at the MIT, is studying a method for transmitting energy25 wirelessly). The demand for mobility is the reason why also virtual worlds should become accessible not only on computer terminals, but also using mobile devices. In order to achieve this, virtual worlds should have multilevel architectures in terms of detail, richness and resolution of the visualized contents and should automatically detect the users’ terminal characteristics and consequently adapt the level of details of contents.

- **Online virtual worlds communicating with each other**: at the moment there are dozens of different virtual worlds available on line; for each one the user has to create a different avatar. According to Gartner Group, the world’s leading information technology research and advisory company, in 2011 80% of people that habitually uses internet will have an avatar26, this means that internet will probably evolve towards the direction of immersive applications, and many website will give the users the possibility to navigate them in an immersive way. If virtual worlds, as someone states, represent the first phase of this evolution of Internet into an immersive and experiential web, may be they will have to involve into a network of virtual on line environments compatible with each other so that the user will be able to easily jump from one to another with her/his unique avatar. The same compatibility should be valid also for the objects created by users that should be easily transferred from a world to another.

- **Secure virtual environments**: Until now the most popular virtual worlds have been represented by MUDs (Multi User Dungeon or Multi User Dimension) and MMORPGs (Massively Multiplayer Online Role-Playing Game); however, nowadays virtual worlds are not only created for entertainment but also for business and education. Many companies are looking at virtual worlds as a new platform for training their employees, testing new products or experimenting marketing strategies. At present virtual worlds can support many real life activities but most of their potentialities are still unexplored. In order to reach a massive use of virtual environments for business purposes, these environments should first of all be supported by secure infrastructures in which important transactions can occur. Thus it is essential to develop systems of identification and authentication and to create a kind of virtual legislation for guaranteeing and protecting the rights of virtual citizens.

---


• **Changes in real world reflecting in virtual world**: The ideal situation would be when the changes in the real world would reflect in the virtual world automatically and immediately. For example, if a car moves in the real world, then a 3D map in the virtual world would update its position automatically or a virtual therapist who would be able to sense the mood of her real patient and provide adequate psychiatric assistance. This is going to require research in many inter-disciplinary domains including networking, computer graphics and real-time computer simulation. As award winning author, Vernor Vinge says in the June 2008 issue of the IEEE Spectrum\(^{27}\), about the pure Internet Scenario - when humanity and its networks and databases become a seamless entity and acquire a collective (possibly superhuman) intelligence:

> "Though physical objects need not be individually sapient, most would know what they are, where they are, and be able to communicate with their neighbors (and so potentially with the world).... In general, I suspect that machine/network life-forms will be faster, more labile, and more varied than what we see in biology."

### 5.6 User centric search & retrieval

Multimedia content, which is available over the internet, is increased in a rate faster than the respective increase of computational power and storage capabilities. Internet capacity will approach the amount of yota \((10^{24})\) bytes in 2010. Such a tremendous amount of content cannot be processed and indexed by the current computational power unless personalised and user-centric mechanisms are implemented so that only the content of interest is delivered to the end-users.

First, search and indexing of multimedia data should be performed in a dynamic and scalable way. Today content is dynamically created, deleted, changed and retransformed in distributed and pervasive environments. Therefore, content indexing methods should be dynamic and autonomous in the sense that they can change their structures as content is evolved. In addition, scalable mechanisms are required for efficiently handling the huge amount of information. Scalability concerns not only the classical view of algorithms, methods and tools used for managing and organizing huge amount of data, but it also reflects the schemas required for handling the really dynamic, evolving and pervasive behavior of the content. In this direction, the conventional searching mechanisms should be extended and modified to tailor with highly unstructured peer to peer networks in which new nodes become active/inactive in a continuously dynamic manner.

Second, we need to pay a great research attention on increasing the content utilization efficiency, measured as the fraction of the relevant delivered content (i.e., content which satisfies their information needs and preferences) over the total amount of delivered content. This goal can be achieved by three main research actions. (i) The first concerns the development of interoperable interfaces able to describe multimedia content under a context aware and user-centric framework. That is, different tags will be derived for different users or different social groups, resulting in personalised content tagging schemes. (ii) The network will be able to filter out all the unnecessary information incorporating context-aware semantic routing mechanisms. (iii) Multimedia content indices are organized in a scalable and personalised, user-centric framework resulting in different structures for different users by exploiting, for example, social-based relevance feedback mechanisms.

Third, the main bottleneck in indexing, searching and retrieving the distributed multimedia content is that the vast amount of information is not actually annotated. Thus, automatic annotation mechanisms are required to be implemented for personalised description of the distributed content. In particular, the highly unstructured and not annotated content should be automatically transformed into meaningful structures, which describe the context-aware information of the content, ranging from content characteristics, network properties and user's preferences.

6 Making an Economic & Social Impact in Europe and Worldwide

6.1 European IPR

6.1.1 Trends in managing copyrights

Copyright law is meant to stimulate innovation by striking a balance between the interests of the author and the public interest. Every work, if it is considered to be an original creative expression, falls automatically under copyright jurisdiction. As markets, devices and services are converging, there is a growing need to enable easy and fast circulation of copyrighted content. As Screendigest concludes, being able to let content circulate at moderate transaction and coordination costs is crucial in getting busi-

ness models for established as well as new players feasible. Several developments cause fierce discussion over the enforcement of copyrights:\(^{29}\):

- First of all, the physical and technological barriers to copy information are eroding. The global scale on which this takes place is also one of the main contemporary problems as the internet is not bound by geographical borders which makes it very difficult to define where lawsuits can be filed and where they will be heard in court, what legal laws should be applied as there are more jurisdictions\(^ {30}\).

- Secondly, the fact that users are taking up new roles in aggregating and sharing content is a development affecting business and therefore the discussion on copyright. A landmark development in this respect was the rise of file sharing platforms such as Napster, KazAA and eMule.

- Another role users are now taking up is the creation and publishing of content themselves. Over 75% of all content is now created outside professional practices. Where family pictures and private videos were once secluded to the home environment, this material is now published on-line for practical purposes like sharing with friends. Most of this material was not created by people with professional intentions. A lot of these user generated photos and videos were never part of the economic system that inspired the design of the copyright law\(^ {31}\).

The challenge for service is to provide a solution for professional content owners in the dilemma between keeping control over the content while at the same time enabling circulation of this content. The urgency for a level playing field between stakeholders here is high since both copyright owners as well as service providers have learned that taking cases of copyright infringement to court is not necessarily in the interest of both sides.

---


6.1.2 Digital rights Management

Keeping control over copyrighted material and leveraging these rights to create value requires some kind of management of rights. DRM is normally associated with publishers and other copyright holders preventing access, copying or conversion of digitized content\textsuperscript{32}. In the light of the growing need of legal circulation of digital content, DRM is growing into an integral part of the total management of content:

As the words ‘digital rights management’ suggest, DRM systems are based on digital technologies that describe and identify content, and enforce rules set by right holders or prescribed by law for the distribution and use of content. […] DRM systems do not impede access or copying per se, but rather create an environment in which various types of use, including copying, are only practically possible in compliance with the terms set by the right holders. Therefore, they usually do not deny access but rather manage access to content by combining technical measures with a payment mechanism.

Management of content can only partially delegated to technology. Checking whether user generated content contains copyright holding material of third parties for example is not a fully automated process, and requires human moderation. The same applies when it comes to DRM systems. Syndication, creating new windows for content or other ways of circulation content might be fully automated, but up until now this proves to limit the possibilities to leverage the rights and create value. Most DRM systems are not yet capable of fulfilling this promise for several reasons.

- A first obstacle is the limited interoperability of DRM systems. As management of digital rights is now organized around the protection of those rights, the systems designed with this perspective in mind proved difficult to combine with other DRM systems, leading to compatibility and usability problems.

• A second obstacle is related to user generated content. Most DRM systems fail to take into account that amateurs are not always interested in exploitation of the material. They were designed with exploitation of the copyrighted material in mind, not taking into account different ways of non-commercial licensing to a service provider.

New developments in the field of software architectures might prove to be a solution in taking away these bottlenecks. Instead of protecting content by embedding the protection in hardware or encrypting the content, Service Oriented Architectures (SOAs) enable the management of rights on digital content by metadata by means of encryption, fingerprinting, tracking software (“where is my content”) and disabling properties. The challenge here is to create disabling and take down capabilities and tracking software that can stand the bypassing tricks of hackers.

If these kinds of systems prove to be robust, they might offer service providers and right owners several advantages:

• Increased flexibility in licensing agreements. Managing rights based on metadata increases the possibilities to tailor the licensing agreement to the specific needs and wishes of the content owner. Especially amateurs without commercial motives might be served in this way, for example by integrating Creative Commons licenses as an option in the terms of use of a service.

• DRM based on metadata also provides the basis for combination with numerous other metadata based business applications. For example, combining data on copyright status of a given content clip with user metrics might prove to be an effective way to manage right clearance and royalties between service provider and content creator, either amateur or professional. A more exotic example is implementing a system that allows for automatic devolution in non use.

Making DRM systems part of a service platform based on such a database might enable service providers to tailor the terms of use to specific needs, to leveraging rights more effectively and make financial arrangements more evidence based. In sum, the challenge for all stakeholders to achieve consensus on standards, protocols and business processes to manage rights as an integral part of their daily business activities.

6.2 An ecosystem for innovation

In the converging marketplace for media, traditional value chains become more networked, crossing the established boundaries between once separately operating creators, producers, packagers and distributors. This leads to the emergence of ecosystems of innovation, in which value is created by a high number of actors (including
users in different roles), contributing most of the time in a non-linear, loosely organized manner. This leads to a high degree of interdependency between actors. Cooperation and competition might take place at the same time. Instead of creating value by managing the different roles separately, value creation is increasingly taking place in building and maintaining the relations between these roles.

Users are increasingly part of this innovation process, transcending their role of consumer and adding value just as any other business actor (Fig. 3).

Figure 3: Users as value adding actors. (Source: Limonard & Esmeijer (2007))

Although users are increasingly acting as “prosumers”, it has become increasingly difficult to reach them. As more and more service providers approach the customer

---

together, and over different channels, “proximity to the customer” might be a more appropriate concept to outline the challenges for companies active in this field. Within the domain of media, cross media strategies are increasingly applied to create this proximity towards the customer. Media companies develop concepts and forge partnerships to optimize the channel mix, creating an immersive experience and applying different revenue models.

This relationship management is not confined to consumer markets. As the relationship to customers gets more dynamic and volatile, suppliers of content also feel a growing need to have the opportunity to cater to their customers in a more flexible way. Service providers therefore not only need to adapt to the changes in their user base, but are also pressured to constantly monitor and adapt to the changing position of their business partners. Developments on the internet and within the ICT industry in general have significantly lowered barriers to exchange information and trade, enabling direct contact between economic agents to such a degree that extensive disintermediation may take place. In other words, bypassing intermediaries has never been easier. The challenge for service providers developing new services is to strike the right balance between profiting from network externalities and keeping control and capitalizing on the assets this actor brings in into this business network. Key assets within the domain of media are control over copyrights on different types of platforms, a (loyal) base of existing customers, brand equity and service platform capabilities.

Although different ways of advertising seem to be the dominant revenue models in media, the range of potential revenue models is growing. Affiliate marketing, third party service provisioning, pay-per-use, different subscription as well as “razor and blade” models of cross-subsidizing are also growing in popularity. The most critical choice here seems to lie in the mix if the different revenue models applied, whether they create synergy and match with the embedded routines of the target group.

**From Mass Media to Custom Media**

The development of web 2.0 has been extremely innovative the entire market. With the user shifting from a passive into an active role, it's clear that also a new figure of consumer is emerging. We can define her/him much more independent and less impressionable by traditional advertisement. The new consumer exactly knows what she/he wants and most of the time knows where to find it. This is the reason why it becomes necessary for companies to diversify not only their offer but also their way of communicating it; traditional advertisement have been replaced by new forms of spots and promotion that are mainly tailored on purpose for the final consumer. Experts speak about an important transition from mass media into custom media.

Customization of marketing strategies also means taking into account brand new forms of communication, such as web 2.0 applications and virtual worlds. At the moment we can already detect many examples of brands that have been promoted by means of these media; may be the first results are not as brilliant as expected, but the
trend is clear. FI will represent the perfect realization of the custom media concept. Thanks to FI companies will be able to reach customers one by one.

**Mass product VS niche products**

Another important and parallel evolution is implied in the development of FI. As perfectly described in the Long Tail theory, elaborated in 2004 by Chris Anderson\(^{34}\), the total demand for niche products (in the figure the long yellow tail) is as big as that for mass ones. The reason why, since a few years ago, mass products dominated the entire market without any, or little, space for niche products, has been due to the fact that distribution was monopolized by majors and large retailers and communication passed through mass media, with prohibitive costs for small producers and retailers.

Now that anyone can have access to virtual cheap displays and ship a product at low costs, niche products can potentially enjoy the same visibility and success of mass product. Here again the role of FMI will be strategic for a complete change of paradigm.

### 6.3 Standardisation

Standardising the mechanisms for 3D media delivery, personalisation, coding, context-awareing, etc. should be contemplated in the plans of the technology developments. Two of the success cases in the European standardisation are Digital TV (e.g. DVB) and in Mobile communications (GSM and beyond). Future Internet research in Europe, is expected to win the market race to other parts of the World, and standardisation has been a critical fact in some of the technologies which will be part anyway in the Future Internet domain.

In the past, the standardization efforts have been made with preference to the physical layers and the network protocols. But is key in a horizontal market like Future Media Internet is expected that all the layers, including the applications and services interfaces could be harmonized and where appropriate standardized.

Not only the standardisation is important for the network protocols but also for the rest of the layers which conform the Future Media Internet the European Commission is promoting.

Just not to overextend this section with several standardisation bodies and fora, the relevant ones are indicated where appropriate in the text.

---

\(^{34}\) Chris Anderson, “The Long Tail”, Hyperion, 2006
Harmonisation of mechanisms and procedures is needed in the field of media applications and services, where standardisation is not the main purpose. Agreements among players should be promoted and only with joint research it will be possible for Europe to take a key role in the Future Internet economy.

7 International collaboration

The opportunities that the development of high speed, pervasive connectivity will bring to International collaborations will empower the way to obtain mutual benefits between Europe and the rest of the world. In fact, in the context of the FI, several complementarities between the strengths of European and other economies are starting to be evident (e.g. European leadership in iDTV and mobile technologies and potential of Asian countries in software development, search engines, etc) and will increase in the future. The opportune and early identification of synergies between Europe and non-European countries will foster the establishment of partnerships with specific countries that give added value in specific technological domains and market opportunities in the context of identification of clear mutual benefits.

More specifically, multiple regional initiatives are currently emerging in view of defining future global networks. Japan (through the AKARI Architecture Design Project\(^{35} \)) and Korea\(^{36} \) have made public their ambitious initiatives, China is supporting the domain through an ambitious and integrated industrial policy, in the US the FIND and GENI programmes\(^{37} \) and facility is a key contributor to the debate on the future of the Internet and with Latin America there are several ongoing initiatives for identification of opportunities for ICT collaborations\(^{38,39} \). These initiatives are not all tackling the issue of the Internet evolution as part of their core objectives, but are certainly related to technological and socio-economic scenarios (ubiquity, connected devices) that will clearly need to be taken into account when addressing the Internet of Tomorrow.

It is clear that Europe should reinforce its position, but what should be firstly strengthen is to have a unique position in the Future Internet research and fight

\(^{35}\) http://akari-project.nict.go.jp/eng/overview.htm


\(^{38}\) http://www.solar-ict.eu/

\(^{39}\) http://www.salamas.eu/
against the effort fragmentation among some European countries, as it has been done with many other European policies in the past.

Europe should work towards playing the key role in Future Media Internet in the World, and should take advantage of the European high-level researchers and the collaborative research they can carry out, as it has been described. European academia and industry cooperation is key and should be strongly promoted with the instruments of the European Commission.

8 Conclusions

This white paper on User Centric Future Media Internet reflects the work that is currently being carried out from the User Centric Media cluster of ongoing FP6 & FP7 EU funded projects. The main objective of this work was to identify the challenges that should be realised in order to reach the goal of the User centric Future Media Internet which is one of the main pillars of the Future Internet.

Towards this aim three working groups were formed within the body of UCM namely: the Personalized & Creative Media WG, the Future Internet of Immersive Media WG and the Future Media Internet Communities WG. After explaining the main concepts which comprise the user-centred idea, the provisioned research challenges from all WGs were reported, such as: content aware networks and network aware content, user generated content, collaboration and interaction, immersive media experiences beyond HDTV, co-existence of virtual with physical worlds and user centric search and retrieval.

The UCM cluster believes that specific research is needed in a) Future Media Internet Networks for providing an active role to the individuals, supporting efficient delivery, retrieval and management of 3D data volumes, understanding how to provide appropriate abstractions content services and how they can form a content network on top of overlay networks, providing appropriate file format for transmitting 3D content and enabling to filter out the transmitted information according to the interests of the end-users; b) harmonising the transmission and signalling of the multimodal content across networks; c) involving the virtual worlds, which represent the first phase of the evolution of Internet, into an immersive and experiential web making the virtual on line environments compatible with each other so that the user will be able to easily jump from one to another with her/his unique identification.

Finally, the white paper concluded with an exhaustive analysis of the impact that these challenges are expected to have in the world and presented other relevant initiatives which are currently taking place worldwide showing the potential for international collaboration.
Annex I - Editors & Contributors

Editors

Isidro Laso-Ballesteros (Cluster Leader)
European Commission
Networked Media Syst. Unit
Isidro.Laso@ec.europa.eu

Petros Daras (UCM cluster Coordinator)
Centre for Research and Technology Hellas - Informatics and Telematics Institute
daras@iti.gr

Contributors

<table>
<thead>
<tr>
<th>WG1: Personalized &amp; Creative Media</th>
<th>WG2: 3D &amp; Immersive Media</th>
<th>WG3: Future Media Internet Communities</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIVAS</td>
<td>VirtualLife</td>
<td>Samba</td>
</tr>
<tr>
<td>Nick Achilleopoulos</td>
<td>Marianna Panebarco</td>
<td>Oscar Mayora, Margarita Anastassova</td>
</tr>
<tr>
<td>Intermedia</td>
<td>VICTORY</td>
<td>InterMedia</td>
</tr>
<tr>
<td>Nadia Magnenat-Thalmann, George Papagiannakis</td>
<td>Petros Daras</td>
<td>Nadia Magnenat-Thalmann, George Papagiannakis</td>
</tr>
<tr>
<td>Samba</td>
<td></td>
<td>Playmancer</td>
</tr>
<tr>
<td>Oscar Mayora, Margarita Anastassova</td>
<td></td>
<td>Elias Kalapanidas</td>
</tr>
<tr>
<td>Content</td>
<td>SAME</td>
<td>Content</td>
</tr>
<tr>
<td>Carmen Guerrero</td>
<td>Antonio Camurri, Gualtier Volpe</td>
<td>Carmen Guerrero</td>
</tr>
<tr>
<td>CitizenMedia</td>
<td></td>
<td>CitizenMedia</td>
</tr>
<tr>
<td>Carmen MacWilliams, Petter Bae Brandtzæg</td>
<td></td>
<td>Carmen MacWilliams, Petter Bae Brandtzæg</td>
</tr>
<tr>
<td>MyMedia</td>
<td></td>
<td>MyMedia</td>
</tr>
<tr>
<td>Alexander Voss</td>
<td></td>
<td>Alexander Voss</td>
</tr>
<tr>
<td>TA2</td>
<td>TA2</td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td></td>
</tr>
<tr>
<td>Peter Stollenmayer, Doug Williams, Ian Kegel, Marian Ursu, Kees Tuinenbreijer, Annika Waern, Rob Willems, Joke Kort, Sander Limonard Manfred Lutzky, Fabian Küch, Oliver Hellmuth</td>
<td>Peter Stollenmayer, Doug Williams, Ian Kegel, Marian Ursu, Kees Tuinenbreijer, Annika Waern, Rob Willems, Joke Kort, Sander Limonard Manfred Lutzky, Fabian Küch, Oliver Hellmuth</td>
<td></td>
</tr>
<tr>
<td>My-e-Director 2012</td>
<td>Games@Large</td>
<td></td>
</tr>
<tr>
<td>Nikolaos Papaoulakis</td>
<td>Alex Shani, Wijnand IJsselsteijn</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VISNET II</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Janco Calic</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CHORUS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Jean-Yves Le Moine</td>
<td></td>
</tr>
</tbody>
</table>