Networked Media of the future

October 2007,

Brussels

NETWORKED MEDIA OF THE FUTURE

The term Networked Media encapsulates the concept that in the future all the elements of the media value chain will have a network capacity attached to them. Not only does it mean that all devices and multimedia content will be network-enabled but it also means that users and providers will participate and collaborate actively in a community network.

The Networked Media Task Force (NM-TF) Vision is built upon three axes that will be available in the future, i.e. true broadband, personalised media, and distributed control. The NM-TF envisions ***"new models of rich media interactions and cooperation based on enhanced AV content at the community level with distributed control infrastructure and community intelligence allowing pervasive personalised media services without the need of central control".***

This vision of pervasive personalised media services without a central control requires research in several domains. In terms of *content and services*, there is a need to enable non-professional production of content, creating services that reflect societal needs and facilitate the emergence of multi-party, multi-play services. In terms of *access*, people require to retrieve their multimedia assets and use them on their devices while being mobile. In terms of *collaboration*, technologies need to leverage the increasing computing power, data capacity and connectivity of surrounding devices, and exploit them in a distributed fashion. It will require as well to enhance *virtual reality and conferencing* technologies to enable a more natural and effective team collaboration, and also *empower virtual communities* to create content and services.

Consequently, research need to focus on creating an environment, where all of the human senses are engaged (but not overwhelmed) and where communication via technological means preserves (and maybe augment) the richness and subtle characteristics of human-to-human communications.

The vision in a nutshell

"New models of rich media interactions and cooperation based on enhanced AV content at the community level with distributed control infrastructure and community intelligence allowing pervasive personalised media services without the

The overall impact foreseen by the NM-TF is enormous modifying the ways citizens and organisations relate among themselves while also affecting the market rules and structure of the media sector. The networking and media communities will end-up integrated so as to provide audio-visual information seamlessly across the entire range of delivery networks to the whole range of possible user platforms and it will restructure the research capacities and technological skills integrating them into a global vision for new networked media environments.

This document produced by the three Networks of Excellence in the area of Networked Media funded under the EC Framework Programme 6 (VISNET-II[1], INTERMEDIA[2] and CONTENT[3]), provides a comprehensive vision on the Networked Media of the future and the long term research challenges involved. It will provide a framework for industry, academia, governments and citizens to act together in order to find the responses to the associated challenges providing socially acceptable solutions.

The EC is committed to continue promoting research in a field that is going to affect everybody in every walk of life, be it for leisure or for access to culture and education, just to mention some examples. This is the case already today and will be even more so in the future as personalised media services without a central control will be pervasive underpinning the media revolution that is in the making.

Luis Rodríguez-Roselló

Head of Unit

"Networked Media Systems" Unit

Directorate General Information Society and Media

European Commission

Isidro Laso Ballesteros

Scientific Officer

"Networked Media Systems" Unit

Directorate General Information Society and Media

European Commission

[1] VISNET-II NoE. Networked Audiovisual Media Technologies.
Coordinator: Prof. Ahmet Kondoz, University of Surrey, UK.
http://www.visnet-noe.org

[2] INTERMEDIA NoE. Interactive Media with Personal Networked Devices.
Coordinator: Prof. Nadia Magnenat-Thalmann, University of Geneva, Switzerland.
http://intermedia.miralab.unige.ch

[3] CONTENT NoE. Content Network and Services for Home Users.
Coordinator: Prof. Arturo Azcorra, University Carlos III Madrid, Spain.
http://www.ist-content.eu

TABLE OF CONTENT

1 EXECUTIVE SUMMARY

The term “Networked Media” implies that *all kinds of media including text, image, 3D graphics, audio and video are produced, distributed, shared, managed and consumed through various networks like the Internet, WiFi, GPRS, 3G and so on, in a convergent manner.* As network and device technologies grow and are improved more with the advent of ubiquitous computing environments, Networked Media have become important not only to daily life but also other fields such as entertainment, business, and production processes.

Therefore, the Networked Media sector needs to start taking into consideration the continuous advancement of key technologies such as Information Technology, Networking, Electronic Equipments and Content production/consumption, all of which are needed to achieve and provide useful Networked Media services. *Our vision here is to increase socio-economic impact by the introduction of technologies that enable: a) a richer media exchange; b) enhanced AV content-based interactions at the Community level; c) distributed control infrastructure; d) Community intelligence, allowing personalization of services without the need of central control.*

In the current state-of-the-art in the field of Networked Media, we investigated available networked media technologies, systems and platforms that cover current mobile communication support for multimedia. We also discussed current status of Networked Media creation, discovery, selection, distribution and consumption issues, as they are required procedures to serve any Networked Media service. They address lots of points that have to be considered including, mobility support in different networks, privacy, quality of services, energy efficiency in mobile devices and multimedia content adaptation. Although, all of these issues are tackled in ongoing research activities such as INTERMEDIA, CONTENT, VISNET II, NEM, NETS and FIND in the current context and situation, it is required to go further when we think about the Networked Media of the future, which will incur more advanced and complicated challenging issues to be solved.

In this white paper, the expected networked media of the future is drawn around three main axes, after several physical meetings and brainstorming sessions by the Networked Media Task Force (NM-TF): true broadband, personalised media and the challenge of distributed control. To discuss these axes, three challenges have been discussed:

- “***The challenge of true broadband***” is a new situation in which Gigabits per second (or even higher) connectivity will be supported to not only in our homes, but also wherever we go/travel and whenever we would like, soon. The real question here is indeed: “*what will we be able to do with this very high-speed connection, which we cannot do now?*” In order to answer such demanding questions, it is firstly necessary to identify the applications that can benefit from such high speeds as well as their associated research challenges. We identified and discussed new research challenges and applications, which affect our environments: virtual collaboration environments, user based new services, service continuity support, heterogeneity, edutainment, virtual sports groups, telepresence, device cooperation, privacy and security and ubiquitous interfaces.

- "***The challenge of personalised media***" is a new situation in networked media. The real question here is indeed: *how recent advances in related networked media disciplines can influence them?* To answer such a demanding question, it was necessary to elaborate on the two most prominent key research directions as identified by the multidisciplinary NM-TF group: 1. Real-time and interactivity and 2. User domain and inclusion. The former research direction deals with issues such as how to compensate memory and computer power with network bandwidth, interactivity support, real-time transmission of large media content and databases, and enhancing realism of virtual human and robotic behaviours. The latter research direction covers topics such as understanding the limits of the applications, cultural diversity in media, user-domain and application inclusion, multi-disciplinary research support and seamless framework.

- "***The challenge of distributed control***" is a new situation in which neither the infrastructure, nor the service, is controlled by a single entity, to deliver a specific service. To address this challenge we discussed possible issues in various directions, from underlying network support to content management including infrastructure requirements, dynamic networking issues for distributing content, content service infrastructure, and content search, summarization and adaptation for efficient distribution.

IMPACT OF THE 'NETWORKED MEDIA OF THE FUTURE'

The overall impact that we foresee is that the new developments in Networked Audiovisual Technologies will fundamentally modify the way citizens and organizations relate among themselves. It will also affect the market rules and structure of the media economical sector, as well as opening huge opportunities for new business sectors.

More in detail:

- In technological areas, research and development (R&D) can potentially make a huge impact on the providers and consumers of new audiovisual products and services, especially in *the increasingly pervasive content networking environment*.

- During the past decade, telecommunications and computer networking have essentially merged their R&D activities, and are also increasingly merging their infrastructures and business sectors. On the other hand, networking and media communities have not joined forces yet, and are still essentially separate research (and industrial) communities. Yet, the trend is unstoppable, and *the networking and media communities will end-up integrated,* whether they like it or not. Incumbents and new-entrants understanding this phenomenon may get huge profits by taking advantage of the opportunities that will arise. Incumbent companies, no matter how large and powerful they are now, which do not believe in this trend, do not understand it, or just do not have the ability to re-invent themselves in this new framework, will be driven out of the business, or acquired by a rival or newcomer that has the ability to adapt.

- With respect to people and economy, Networked Media *will enable natural communication between people residing in distant locations* without incurring the cost and trouble of travelling. Economically, the use of Networked Media will result in direct cost savings, and more productivity for dispersed teams. The access to educational material via the network and presenting it in media-rich format, will enrich the learning experience, and improve the understanding of the learners. The ability to access the sheer amount of data stored in different computing devices at home, including pictures, videos and music, will make personal data truly ubiquitous. This will radically change people's consumption patterns of multimedia content with one multi-purpose device instead of multiple single-application devices. The provision of cheaper health monitoring systems will extend medical care in Europe to people living in remote areas, the aged population and citizens with physical disabilities, from the comfort of their homes.

- From the value-chain point of view, one of the main impacts on the networked media research is *dissolving technical boundaries of different disciplines* related to the personalized networked media environment. Networked Media will also have a strong impact on restructuring the research capacities and technological skills in the wireless wearable mobile media field. In addition, Networked Media research will provide international research infrastructures for research activities in the field of wireless wearable mobile media, which will be developed and maintained for the needs of other communities as well. Networked Media will generate a new vision on the nomadic daily life as the development of networked media devices will take new standardization from standalone mechanical devices to networking and adaptable devices.

CHALLENGES AHEAD

In order to meet the above challenging goals, the future Networked Media should be designed, developed and released with the following considerations:

- *New form of the Networked Media*, which includes large amount of meta-data and user generated tags so that a system can easily analyze content in a semantic way.
- *Easy way to access the Networked Media*, which will be distributed in different geographic repositories and connected to the users with their heterogeneous (mobile or stationary) devices in various underlying network conditions.
- *Protection of the Networked Media from any illegal and malicious usage*, since all content can be accessed anywhere and at any time, it is essential to provide secure and safe access to private and secret information, lest it should be susceptible to illegal use.
- *Payment models for commercialization of the Networked Media in new business markets*, which will become more complicated than those in the current situation because the Networked Media will be produced, managed and consumed by tightly linked interaction among various kinds of service providers, producers and end-users.

2 VISION

To maintain Europe's leading role in the service and technology arena and increase the benefits to its population and economy, further development of the core technologies need to be facilitated. In the age of multi-aspect convergences and personalisation of services, the required developments need to take place in the different enabling technologies (networking, content, interaction), but also across them, to enable a seamless integration, and efficient delivery of added-value services to the end-users.

Our vision is that new models of interaction and cooperation will multiply its penetration and social impact because of the introduction of technologies enabling: a) a richer media exchange; b) enhanced AV content-based interactions at the Community level; c) distributed control infrastructure; d) Community intelligence, allowing personalization of services without the need of central control. These will lead to the establishment of Distributed Content Service Infrastructures, as complex service overlays over the commodity IP network. This will support all sorts of B2B and B2C AV interaction, and in particular, will impact the generation of AV content by allowing communities of users the production of specific content/services for them, both in the context of the “long tail”, or applied to (re-purposed) assets created by incumbent content producers.

The Network Media (NM) sector depends on the continuous advancement of key technologies such as Information Technology, Networking, Electronic Equipment and Content. The continual increase in the adoption of these technologies at home by users, illustrated by the growing sale of consumer equipment and broadband Internet, added to the greater appetite for multimedia content, and their increased mobility (supported by the availability of wireless connectivity and enhanced end user devices), have made it critical to continue developing these technologies, anticipating users' needs and exploring new venues for research and development. Economically, these developments will enhance the productivity of the current pool of experts and R&D activities in Europe, stimulate the creation of new niche markets and drive innovation in the Networked Media sector, especially in terms of new services and enhanced access.

The digitalisation of content, and the availability of affordable broadband access and storage, results in the wide adoption of new technology, and fosters the creation of services. Shortening the time-to-market of services and equipment makes consumers expect shorter cycles in the introduction of technological innovations. Besides the logistical and economical efforts needed to deliver the service/device on time, the two bigger challenges relate to the risk due to the immaturity of technology, and the difficulty of wrapping the technology in an easy-to-use package. This requires more innovation (in technology and design) and more foresight when it comes to investigating the practical applications of cutting edge technologies, and anticipating how to best deliver them to the consumers.

It is noticeable that most current technologies are, by their design and purpose, catered for the young generation. However, with the growing proportion of senior citizens in Europe, there is a pressing need to adapt them to the elderly population, as they have greater potential benefits for their well-being and life-style. In contrast with previous aged generations, the current near-retirement population have grown used to technology, either at home, office, car, etc., therefore they have been accustomed to technology, and its impact (social, economical and ecological). Consequently, it is expected that they would have much higher expectations in terms of quality, functionality and sustainability.

Designing devices and services which will appeal to different user populations and cater for distinct usage environments means that **adaptation needs to be built into the system**. In the coming age of ambient intelligence, context awareness will become a necessity rather than a nice feature to have. A "one-size-fit-all" approach will not resolve the problems of heterogeneity of the users, devices, services and environments. For this reason, customisation at the design level is necessary. For services and multimedia content, for instance, **user will be able to select and mix service/content component, and even create their own** (provided there are tools for that purpose). Furthermore, the ability for a multitude of devices to play content and deliver services requires open interfaces and standards for communications and rendering.

Europe has been at the forefront of innovation in different segments of NM research and development, and leading the way in standardisation activities. Noticeable contributions include multimedia compression, multimedia broadcasting and wireless communications. These characteristics of European research should remain at the core of the approach despite "the increasing industrial competition in incompatible offerings". Interoperability can be achieved by demonstrating the macroeconomic benefits of agreeing on common standards and reducing the costs of incompatibility .

The strong positioning of Europe in NM's different enabling technologies and industries, from communication and multimedia broadcasting to content and consumer devices, makes it an ideal place to lead the realisation of the convergence vision. Europe has already embarked on the path to integrate the different developments in communication technologies, compression techniques and broadcasting, and European citizens are starting to experience the early stages of convergence of web streaming, traditional TV broadcasting and 3G-based delivery of content. In the future, more focus will be put on integrating these technologies, extending their reach (by lowering cost and increasing accessibility), while minimising their impact on the environment and the users' health and privacy. The potential to integrate existing and future technologies and offer choices to customers will open up unprecedented opportunities for the creation of services which were inconceivable or very difficult to implement few years ago.

Building up on the three pillars identified in the NEM Vision 2007 (i.e. broadband, convergence, and home and extended home technologies), the future directions will include promising research areas that are currently under heavy development, and they are expected to revolutionise the way we communicate, interact and consume services. In terms of content and services, there is a shift towards enabling non-professional production of content (especially on the Web), creating services that reflect to societal needs, and emergence of multi-party, multi-play services. In terms of access, people are expected to retrieve their multimedia assets and play them on their devices while being mobile. In terms of collaboration, technologies need to leverage the increasing computing power, data capacity and connectivity of surrounding devices, and exploit a distributed fashion. Besides, virtual reality and conferencing facilities will enable a more natural and effective team collaboration in the workplace, and also empower virtual communities to create user-generated content and services. Finally, education and health care are extremely important and beneficial for the European population. Home based e-health and monitoring services can give improvements in well-being and safety, especially for the elderly, people with long-term chronic health conditions, people with disabilities and disadvantaged groups. **Edutainment** encompasses all of the technological aspects described above, and will have a greater impact, and wider target population. By combining virtual reality, high speed connectivity, adaptive and scalable multimedia content, it will be possible to deliver teaching materials with unprecedented level of richness, interactivity and quality.

The Web has re-emerged as a powerful medium for personal and community expression, via the distribution of user-created content. The lower cost of Internet connection, widespread availability of Internet-capable devices and increased power of software tools, have encouraged more people to join in by posting photos, videos, comments, and creating very large virtual networks of friends. Europe is already investing in the development of new web technologies, and increases the penetration of Internet access, but more effort is needed to maintain the innovation edge with respect to the USA and South-East Asia counterparts, who have been, so far, leading this new wave of the Web. Besides the technical challenges to be tackled, it is important to develop these technologies in-line with societal needs of European citizens, to harness self-expression, ease access to public services and education, and celebrate the cultural diversity of Europe.

The digitalisation of human knowledge and memories resulted in a tremendous increase of access to digital assets. These assets cannot be evaluated based solely on their economical or monetary worth, but they - in many cases- carry a personal value that is hard to measure. These bits of data are collected throughout the lives of people and communities, and are generally dispersed across different computing and storage devices. The ability of users to safely store their data, access and organise them in a more personal manner, will enable a more efficient management of the digital assets. Furthermore, the ability to intelligently search and present a variety of content formats will provide the users with more interactive media experiences. Another technical challenge consists in enabling a portable, yet compelling, version of those experiences in mobile conditions.

The expansion of the European Union has resulted in more opportunities for collaboration across the different countries. Virtual collaboration technologies, such as video conferencing, virtual office, etc., can play a great role in harnessing team working. For them to be widely adopted in industry and academia, as well as in community organisations, they have to demonstrate a clear advantage over face-to-face meetings, or at least, show complementarities with existing approaches. While the cost and ecological impact of virtual collaboration is obvious, their potential advantages of convenience and practicality are less apparent. Consequently, research need to focus on creating a sophisticated environment, where all of the human senses are engaged (but not overwhelmed) and where communication via technological means preserves (and maybe augment) the richness and subtle characteristics of human-to-human communications. As a related application of virtual collaboration, health-oriented services can be provided for the elderly, people suffering from chronic disease and people with disabilities. When combined with sensors placed on the patient, and around his/her premises, they can deliver better services for the most vulnerable people in society, reduce the work load on medical staff and ensure more convenience and dignity the patients.

Virtual technologies can also be exploited for education and entertainment purposes. As people are more eager to self-improve, and have less time allocated for formal educations, the distinction between education and entertainment will become more blurred (edutainment). These technologies will provide a richer learning environment where learners and teachers are engaged in the same virtual space, and have access to material and simulations in support of the courses. Furthermore, virtual technologies can be used for training, assessment and leisure activities (e.g. sport).

It is a significant fact how asynchronous communication forms - like Instant Messaging and Blogging - have reached wide success, yet by offering support only to traditional, text-based information exchange. These two services are the basis of what has been defined as Social Networking, i.e. the possibility for individuals to interact, share, and cooperate with others via the telecommunications technology. The impact of such services can be measured not only in terms of the specific and direct service provided to their user constituency, but also in terms of their contribution to the emergence of new communication styles and language patterns. Particularly interesting are those Social Networks where the link among people is enriched by links to specific places and locations, according to the people-to-people-to-places model (P3).

The availability of **pervasive and affordable mobile networking services** is at the basis of another social communication phenomenon: **Moblogging**. In this communication paradigm people use their mobile terminals to update in real-time a web-diary or Blog, with the insertion of photo, videos and any other audio-visual content. Several commercial services are today available around the world, attracting millions of users that aggregate in forms of quite dynamic virtual communities. A similar model is what has made YouTube service so successful.

More are also emerging. For example, in Rome people are actively participating in a low local TV channel that is airing documents and reports related to the everyday life of the Corviale area. It is important to highlight that most of the AV content is contributed by the people themselves.

3 RATIONALE FOR ACTION

Following the IST 2006 Conference in Helsinki, the DG INFSO Networked Media systems concertation meetings and the recommendations of the NEM initiative, the "Long Term Networked Media Research Task Force" has been created by the coordinators of the **INTERMEDIA, VISNET II and CONTENT NoEs**, under the auspices of the **Networked Media Unit** of the **Directorate General Information Society and Media** of the **European Commission**. Key researchers in this area have been invited and contributed to several physical meetings and brainstorming sessions on the future of Networked Media Research as well as to the current report. In Section 2 of the report, a brief overview of the state-of-the-art in available Networked Media technologies is presented, before the key future questions are addressed in Section 3 and their impact is analyzed in Section 4.

4 AVAILABLE NETWORKED MEDIA TECHNOLOGIES, SYSTEMS AND PLATFORMS

This section provides a concise overview of the state-of-the-art in available Networked Media technologies, systems and platforms, as described in [1][2] (see **Figure 1**). Based on the state-of-the-art discussions, a number of relevant research and technology fields are identified: human-computer interaction and devices, dynamic networking in home environment, management of content, contextual information, software, security issues, and services.

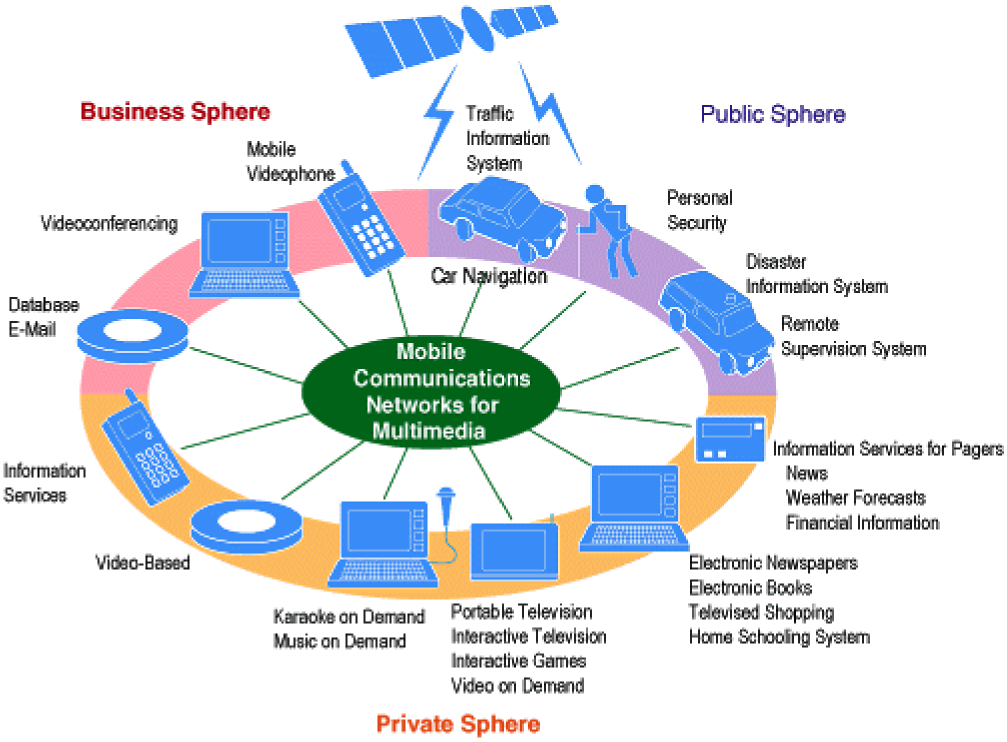


Figure 1: Available mobile communication networks for multimedia (courtesy http://www.cdt.luth.se/~peppar/kurs/smd074/seminars/2/1/2/mpt.html)

In the area of human-computer interaction and devices, multimodal user interfaces are investigated. Speech recognition architectures with a limitation to discrete words or text-to-speech systems are already available, whereas natural speech conversation with computers has not been realized yet. Other possibilities of human-computer communication are kinaesthesia (e.g., gestures), touch (pen-based interfaces or touch screens), tangible user interfaces (combination of digital information and physical artefacts), or emotional user interaction (still in its early stages). Personalized and wearable interfaces as well as seamless interaction of devices are also important in this context.

Communication issues in wireless networks are discussed in the field of dynamic networking in home environment. Wireless and sensor networks are considered in terms of existing standards, QoS requirements, energy efficiency, and mobility. Additionally, an overview and comparison of Bluetooth, Bluetooth2, IEEE 802.15 wireless standards, and ZigBee are presented. This research field also comprises localization and tracking with GPS, GSM, and UMTS, as well as optical, infrared-based and ultrasound-based systems and radio signals. Furthermore, interoperability issues and handoff approaches in homogeneous and heterogeneous networks are presented.

The content field deals with annotation and description of content, content delivery (including end-to-end QoS issues in the IP layer), and content selection (e.g., real-time selection of the most appropriate resource depending on the usage environment). Other important issues concern the management of content and metadata with a focus on MPEG-7 and MPEG-21. Additionally, a state-of-the-art in content adaptation is presented, (e.g., adaptation of scalable virtual characters and the adaptation of the level-of-detail of 3D meshes as well as adaptation of 2D multimedia content in general etc).

In order to adapt user-centric services to changes in the usage environment, the investigation of context-aware systems is necessary. Therefore, the state-of-the-art report comprises different context acquisition and integration techniques and projects, and their relevance to Networked Media. Regarding personalized information management, an overview and a discussion of different approaches for user-modelling as well as a method for ubiquitous user-modelling in ad-hoc networks are presented. The software area deals with middleware, common interfaces, requirements and restrictions for operating systems and protocol stacks. Security is another important research field to be considered in a user-centric multimedia environment. The state-of-the-art report provides issues about Digital Rights Management, privacy, authentication and authorization in different projects like TIRAMISU, ECRYPT, INDICARE, OPERA, and DRMSolutionNG. Additionally, an overview of the commercial DRM systems Helix and Windows Media DRM (WMDRM) are presented. In particular, WMDRM provides real-time encryption of the content, individualization and different licensing terms, which might be of interest for Networked Media. The CORAL consortium and NEMO aim to provide interoperability between different DRM systems through trusted service interfaces. Finally, in matters of services, the service discovery protocols SLP (Service Location Protocol), Jini (Java RMI based), SSDP (Simple Service Discovery Protocol), and UPnP (Universal Plug and Play) are presented. These protocols are used to enable devices to join automatically to a network and find available services.

4.1 Networked Media Content Creation

The annotation of multimedia content spans a manifold set of possible activities which produce descriptions of content on different levels. The so created metadata is supposed to support a huge variety of operations on the content [3]. These metadata could also be **user generated tags**. The task of adapting content to certain requirements, benefits heavily from available metadata. Decisions about useful adaptive manipulations can be based upon the available metadata about content. While such operations will often be related to media format properties, the view on adaptation can be extended to a more complex activity involving more sophisticated personalization steps. Examples of such steps would be the selection or removal of parts of the content based upon e.g., semantic classification or certain depicted entities and situations.

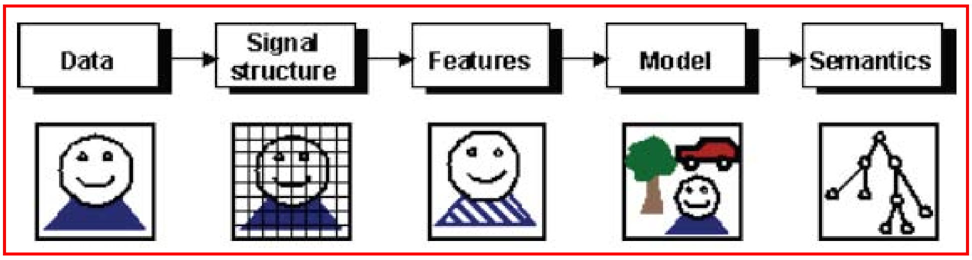


Figure 2: Levels of multimedia content description [1]

Finding multimedia data by certain higher level features or by its semantic content that can be perceived by humans is another important task to be enabled by metadata. Often, search based upon metadata involves a variety of similarity comparisons. Text-based search is applicable if textual metadata, e.g. keywords describing the content, was assigned during annotation.

The metadata produced during annotation can be represented in various ways. Different languages allow for different expressiveness of annotation results and bear different advantages and disadvantages for single tasks, thus specifically developed representations may sometimes be the best choice for a single kind of annotation and a certain usage of the results. MPEG-7 covers many different kinds of annotations and allows for keeping track of results for various further usage scenarios. We can roughly make a distinction between low-level (signal level), medium-level (structural level) and high-level (semantic level) annotations.

Annotations can be performed either by humans or by machines; semi-automatic hybrid procedures with varying degrees of automation are possible. Generally speaking, reliability of automatic annotations is harder to achieve on higher levels than on lower levels. The aspect of the intended use of annotation results is important to keep in mind when deciding for or against certain annotation techniques. In image retrieval for instance, there are various possibilities for sets of features that may serve as a foundation for similarity searches. Depending on what is known about the images to be searched, some efficiently computable feature sets may lead to good results, while they may yield worse comparison results (i.e., results that differ more from how a human would judge) in a different context. As an example, in [4] a suggested approach for medical images and for character recognition involves the comparison of pixel values of images scaled to a certain size, which might not be a satisfying procedure for arbitrary images.

4.2 Networked Media Content Discovery and Selection

Mobile users use their devices in multiple different locations probably including different network setups. Traditionally, network configuration is a tedious and especially manual task that spans all network layers of the OSI-ISO model, from physical layer to application layer. To be fully functional, a networked device does not only need local network access, but needs to have an assigned IP-address for its current subnet, correct router and name server entries. To be accessible for a mobile user, all available application-layer services like printers, file servers, web- and multimedia servers as well as other users of the current subnet have to be detected and presented to the user.

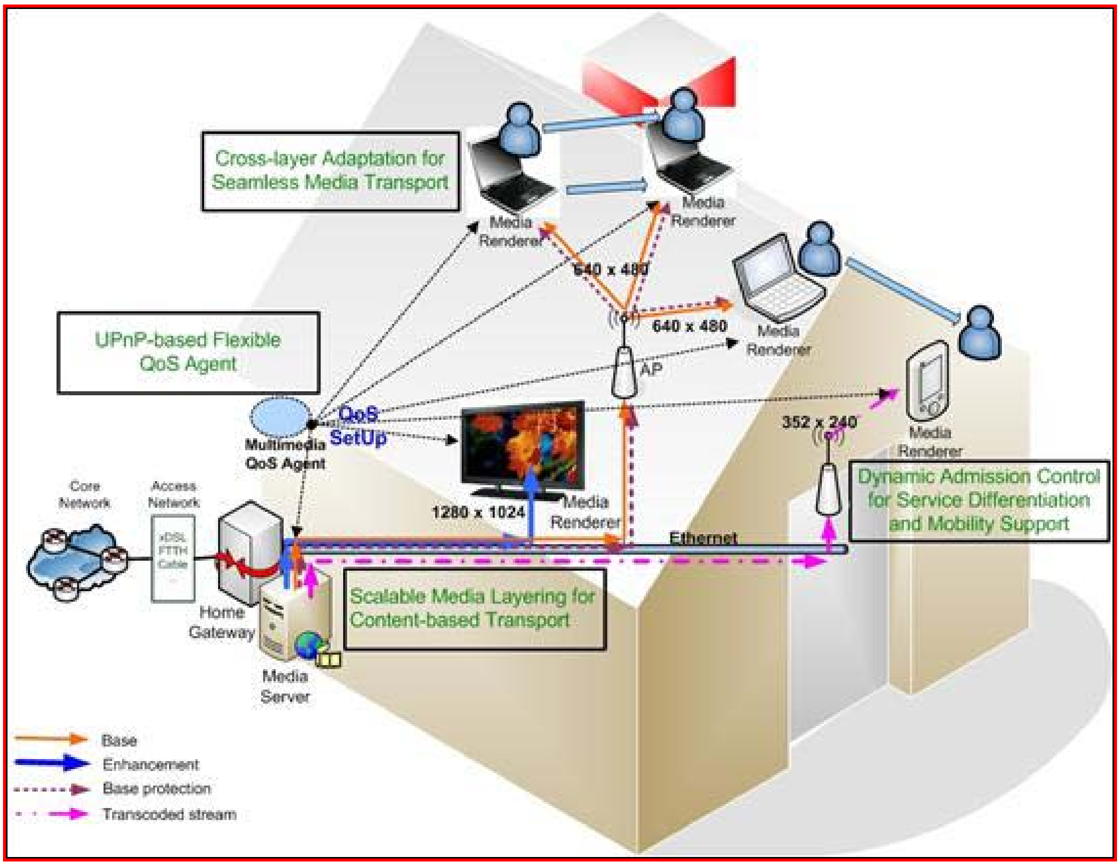


Figure 3: A UPnP home network (courtesy http://p2phome.netmedia.gist.ac.kr/upnp_qos_project.html)

This is a dynamic process because whenever a resource or a service is added to a local network, all users on the network will have to be notified of this. The desired functionalities of dynamic, automatic network setup have been available with early proprietary networking protocols like AppleTalk, Microsoft NETBIOS or Novell IPX, but did not survive the transition to the Internet Protocol (IP). However, a couple of attempts were undertaken to add dynamic networking to IP, most notably SLP, UPnP (see **Figure 3**) and Zeroconf. Nevertheless, all of these protocols do not address the whole chain, from physical access to application layer, and some of the protocols require the setup of a central directory per local network, making them unusable for ad-hoc networks.

4.3 Networked Media Content Distribution

4.3.1 Wireless Networks: Overview and Standards

Today, the most exciting advances in the access network field concern wireless networking. Strictly speaking, any technology that does not use any wired medium could be called wireless networking. However, in the general meaning, they are identified as data packet switched networks that can be categorized according to the size of their coverage area into:

1. Wireless Wide Area Networks (WWAN), which extend over more than a few hundreds of meters, sometimes referred also as Wireless Metropolitan Area Networks (WMAN), while covering an entire town;
2. Wireless Local Area Networks (WLAN), with size between a few meters and a few hundreds of meters;
3. Wireless Personal Area Networks (WPAN), which extend over less than a few meters.

In the latter class, traditionally falls also Wireless Body Area Networks (WBANs), whose size is limited to the human body, and Sensor Networks (SN), whose nodes are typically very simple devices, with hardware for the measurement of several physical quantities, and with limited computation/networking capabilities. Despite the large number of studies in the recent years, the standardization efforts and the market selection have brought about a very limited number of architectures, mostly coming from the IEEE organization: 802.11 for WLAN, Bluetooth/802.15 for WPAN/WBAN, 802.15.4 for SN, and 802.16 for WWAN/WMAN. Many other wireless technologies are available nowadays, such as cellular phone networks (e.g., GSM, GPRS, TETRA, UMTS etc) and digital broadcasting (e.g., satellite communications, DVB-T/S/H etc), but they are not covered here as they exceed the scope of this study.

4.3.2 Quality of Service

Quality of Service (QoS) is a generic term used in communications networks to indicate methods and infrastructures for granting a given service level to selected traffic. Service level parameters may be bandwidth, maximum delay, maximum jitter (i.e., delay variation), maximum packet loss rate (or bit error rate, for networks that allow delivery of corrupted packets). The mentioned service level parameters are deterministic guarantees; statistical guarantees involve statistical service levels, such as mean delay or probability of obtaining a given bandwidth. Some guarantees may be relative, for example a system with two classes of traffic may grant that a high priority traffic class will never get less bandwidth than the low priority traffic.

Several mechanisms can be used to achieve QoS: classification, scheduling, queuing discipline, shaping and policing; they can be used alone or in combination to assure specific QoS guarantees. A classifier is a module that identifies traffic incoming to a network and conceptually tags each packet so that it can be given the agreed QoS. Generally speaking, when an application requests some guarantee, it is requested to comply with a given traffic characterisation. For example, an application requesting a hard maximum for the end-to-end delay should declare what the maximum bandwidth load it produces on a short time basis (peak rate) and on a long time basis (long term rate). A scheduler is a module that decides which packets of those queued should be sent and when. A shaper is a module that, using a scheduler, forces some specifications (e.g., bandwidth limits) on a traffic flow, by delaying or dropping some packets.

4.3.3 Energy Efficiency

Mobile computers typically have limited energy for computing and communications because of the short battery lifetimes. Conserving battery energy in mobiles should be a crucial consideration in designing protocols for mobile computing. In [5], the power drained by the network interface in hand-held devices was studied. An energy efficient probing scheme for error control in link layer is proposed in [6]. The interaction of error control and forward error correction schemes in the link layer are studied from energy efficiency perspective in [7]. The problem of how a mobile in a wireless network should adjust its transmitter power so that the energy consumption is minimized has been considered in [8]. The energy efficiency of a class of multiple access schemes based on a general analytical framework has been studied in [9]. Power aware MAC protocols for wireless ad hoc networks have been studied in [10]. A case for using new power-aware metrics for determining routes in wireless ad-hoc networks is presented in [11]. A collection of papers on energy efficiency in wireless networks is available in [12].

4.3.4 Mobility

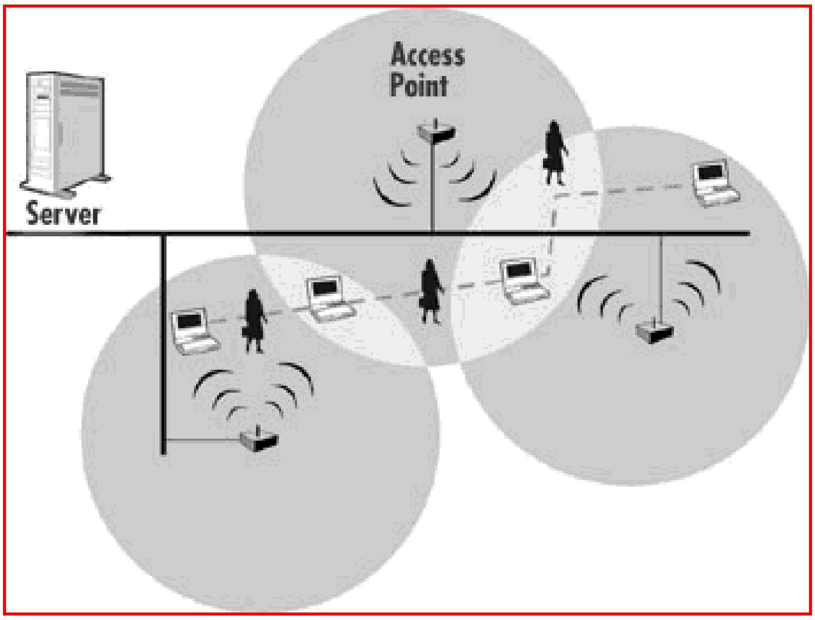


Figure 4: Handing off between WAN access points (courtesy http://www.pulsewan.com)

Handoff (or handover) is the process to assure seamless connectivity for mobile stations that are not communicating in a direct way; it involves the capability of the equipment to switch direct connections between different peers without loosing the connectivity to the third party while re-directing traffic. The handoff procedure (see **Figure 4**) has been traditionally implemented in cellular networks to maintain connectivity of cellular phones to the fixed infrastructure. Among the issues concerning this kind of process, the most significant one is certainly the swiftness: the mobile station must switch from one peer to another one very quickly, otherwise the traffic flow can be seriously affected by abrupt decrease in quality of service perceived, which can lead even to the interruption of the communication. Handover is particularly critical when dealing with voice and video transmissions in data networks. The presence of several protocols in the Internet stacks implies that handovers may be treated at different levels: for example, a mobile station moving between two different Access Points in an IEEE 802.11 network, must change its connection at the data link layer (because of changing Access Point), but it could always change its IP address configuration, if the subnet changes. All of these issues make the handover process more difficult to deal with. Several approaches have been taken for handovers in wireless data networks to date, mainly differentiating for the logical layer at which the handover is performed. Obviously, at the data link layer handover must always be treated, as there is a change in physical connectivity. At the upper layers, when the handover implies a change in the IP subnet, the process can be performed:

- at the network layer (e.g.: Mobile IP [13], Cellular-IP [14], HAWAII [15]);
- at the transport layer (e.g.: TCP-Migrate [16], MSOCKS (TCP Splice) [17], SCTP [18]);
- at the application layer (e.g.: SIP [19]).

Dealing with handover at the network layer is the most common solution, because it works for all applications and transport protocols. However, a proposal to perform handover at the transport and application layers is also interesting, especially when intermediate systems, such as proxies and firewalls, are present, and to take into account the end-to-end Quality of Service degradation that may occur [20].

4.4 Networked Media Content Consumption: Privacy

Today, more than ever, there is an oversupply of different types of computing devices available to the end-user, most of them having high processing power, storage capabilities and capability to connect to various types of networks through different communication interfaces, e.g., wireless and Bluetooth, GSM, UMTS, WiMax, Internet, etc. All of these, combined together, give the user the ability to perform most of his/her everyday tasks anytime, anywhere, and costless over possibly unsafe communication channels like the Internet. This results in taking a lot of security risks that in most of the times are not straightforward visible to him/her, consequently leading to undesirable situations, i.e., user's privacy breaches, unauthorized usage of personal data, etc. As it is well understood, security is of major importance in modern computing systems and special attention has to be given in the deployment and implementation of efficient security mechanisms while designing a system. Special care should be given in the case which we are intended to design a secure embedded system like a PDA, a smart phone with high computing capabilities, etc. [21].

In [1], a short and high-level description of basic security requirements such as authentication, authorization and privacy is provided while references to other security mechanisms like cryptography, key management, hashing techniques, and trusted systems are also provided. These mechanisms are necessary requirements for efficient design and deployment of secure digital rights management (DRM) architectures.

4.5 Ongoing leading edge research in the field

The following projects illustrate the state of the art in the area of large-scale initiatives for networked media:

- **INTERMEDIA:** (Interactive Media with Personal Networked Devices) There have been considerable efforts to have Audio Video systems and applications converge, in particular in home environments with homes as spaces of convergence, and for nomadic users with advanced mobile devices as points of convergence. These trends are important but also have limitations that we seek to address and overcome: home-centric systems fail to account for increased mobility and the desire to provide continuous service across spatial boundaries outside the home; device-centric convergence, e.g. in 3G phones, supports nomadic use but provides a very limited user experience as no single device and interface will fit many different applications well. INTERMEDIA Network of Excellence seeks to progress beyond home and device-centric convergence toward truly user-centric convergence of multimedia. The project vision is ***The User as Multimedia Central***: the user as the point at which services (multimedia applications) and the means for interacting with them (devices and interfaces) converge. Key to this vision is that users are provided with a personalized interface and with personalized content independently of the particular set of physical devices they have available for interaction (on the body, or in their environment), and independently of the physical space in which they are situated. The INTERMEDIA approach to this vision is to investigate a flexible wearable platform that supports dynamic composition of wearable devices, an ad-hoc connection to devices in the environment, a continuous access to multimedia networks, as well as adaptation of content to devices and user context.

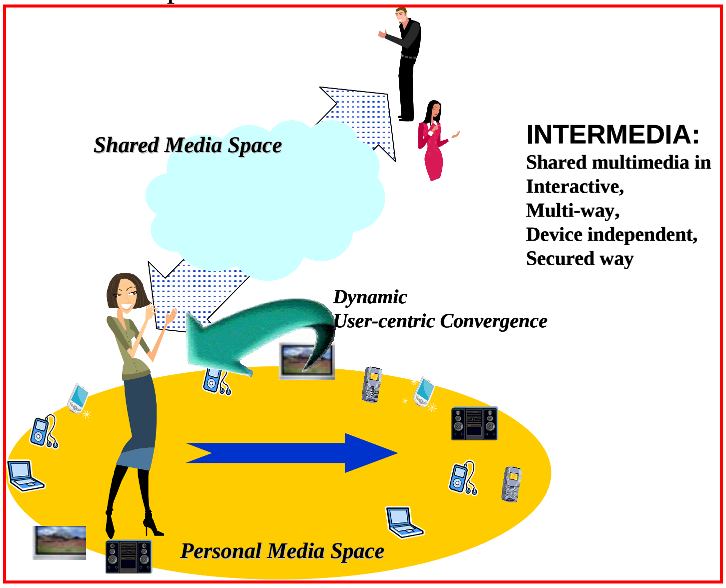


Figure 5: Conceptual Illustration of INTERMEDIA

(Website: http://intermedia.miralab.unige.ch)

- **CONTENT:** (Content Networks and Services for Home Users): The CONTENT Network of Excellence targets a key area of Information Society Technologies, namely Content Delivery Networks for Home Users, as an integral part of Networked Audio-Visual Systems and Home Platforms. CONTENT aims to build the European Research Area in this important communication topic by integrating a group of experts with the purpose of taking forward the state of the art and increasing European leadership in Content Networks. The overall goal of the CONTENT Network-of-Excellence is to integrate the research efforts of the members to address the technical challenges at the different system levels to enable easy-to-install and easy-to-use AV services in and between homes. In particular, the main technical objective will be to boost the potential of European Community Networking by improving Content Distribution infrastructures for the delivery of live (streaming) content and interactive stored content, and by integrating, in an open way, tools and mechanisms that would enable the handling of multimedia assets and their subsequent access for the benefit of the communities of users, producing a set of appropriate services for them, both in the context of the “long tail” or applied to (re-purposed) assets created by traditional broadcasters. CONTENT seeks to integrate the research expertise provided by the project partners on the different levels of (overlay) networks and services for AV content and put special emphasis on the new challenges that arise from community networks and future end-user services for AV content. This leads to three system planes that need to be addressed concurrently: (a) community networks: covering the home networking infrastructure and the access network; (b) overlay networks: proving and end-to-end view for AV services, and (c) content service networks providing a set of services and the composition of simple services to more complex ones. Furthermore, the project addresses the fundamental tradeoffs between non-functional requirements, like efficiency vs. resilience within the planes and between the planes. This requires a well designed coordination between the system planes. In parallel, the consortium performs exchange, education and training of personnel inside and outside the network, dissemination of research results and, in general, spreading of excellence.

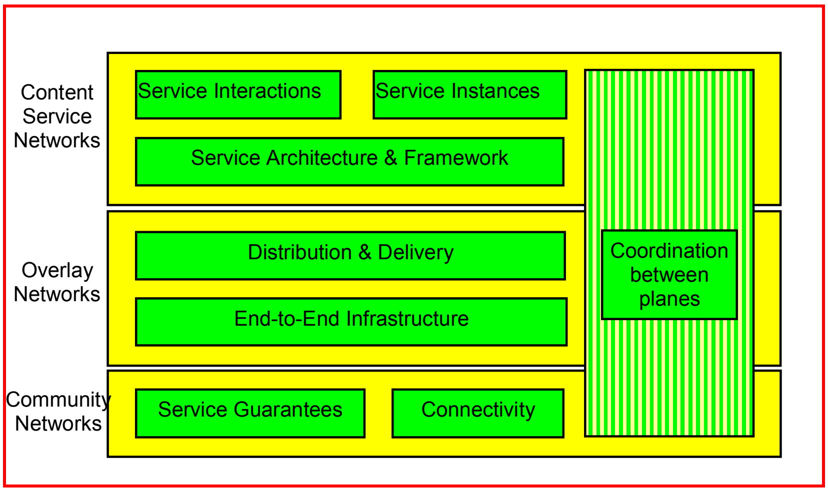


Figure 6: CONTENT Architectural framework

(Website: http://www.ist-content.eu/)

- **VISNET II:** (Networked Audiovisual Media Technologies): VISNET II builds on the success and achievements of the VISNET I NoE to continue the progress towards achieving the NoE mission of creating a sustainable world force in Networked Audiovisual (AV) Media Technologies. VISNET II is a network of excellence with a clear vision for integration, research and dissemination plans. The research activities within VISNET II cover 3 major thematic areas related to networked 2D/3D AV systems and home platforms. These are: a) Video Coding, b) Audiovisual Media Processing, c) Security, as depicted in **Figure 7**. VISNET II brings together 12 leading European organisations in the field of Networked Audiovisual Media Technologies. The consortium consists of organisations known for their proven track record as well as both national and international reputations in audiovisual information technologies. VISNET II integrates a number of researchers who have made significant contributions to the advance of this field of technology through standardisation activities, international publications, conferences and workshops activities, patents as well as many other prestigious achievements. The 12 integrated organisations represent 7 European states spanning across a major part of Europe, thereby promising the efficient dissemination of resulting technological development and exploitation to larger communities.

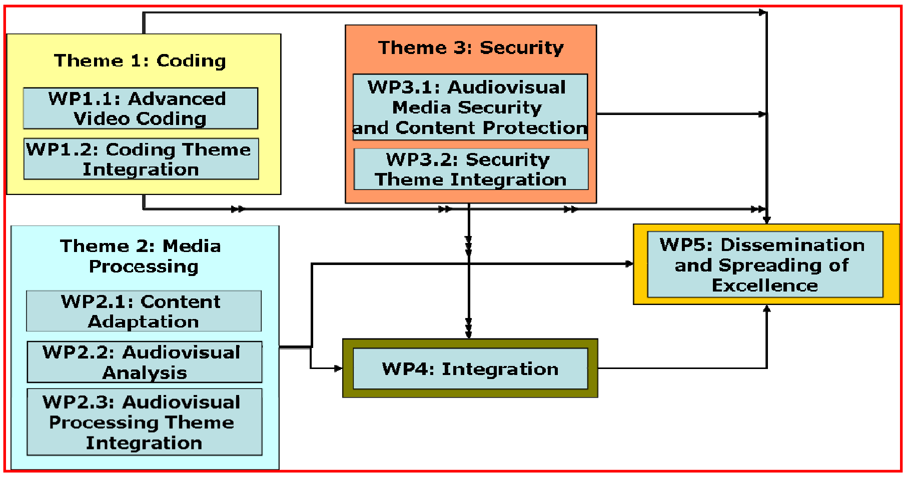


Figure 7: VISNET II thematic areas and their interrelations

(Website: http://www.visnet-noe.org/)

- **NEM:** The Networked and Electronic Media (NEM) Initiative is focused on an innovative mix of various media forms, delivered seamlessly over technologically transparent networks, to improve the quality, enjoyment and value of life. NEM represents the convergence of existing and new technologies, including broadband, mobile and new media across all ICT sectors, to create a new and exciting era of advanced personalised services. The NEM is an industry-led Initiative to promote and direct the large-scale initiative needed to accelerate the pace of innovation and rate of technology evolution to the level that will place European Industry at the forefront of the technology and give users an incredible choice of services. All these efforts will bear in mind the evolutionary framework from home and office environments towards broadband extended home and office environments.

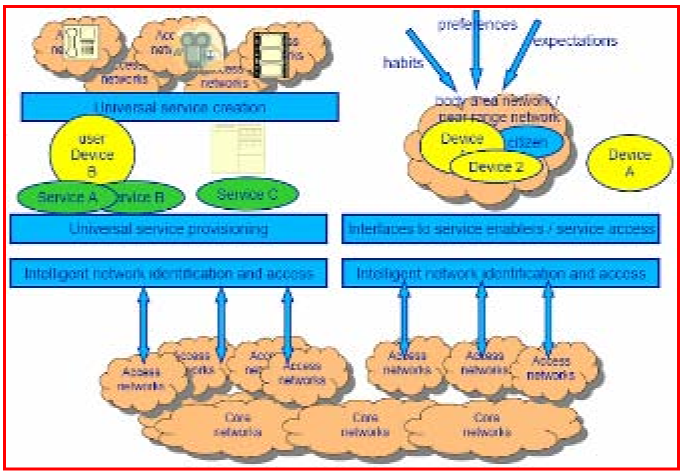


Figure 8: The NEM Initiative

(Website: http://www.nem-initiative.org)

- **NETS:** NSF's Networking Technology and Systems (NeTS) program solicits proposals from the networking research and education community, encouraging investigators to make bold assumptions about the future of networking. The scope of the program covers all properties of information networks including network architecture, protocols, algorithms, and proof of concept implementation of hardware and software. Research activities include creation of new network architectures, modeling of phenomena, network design, analysis, measurement, and performance evaluation. The research scope of the program spans many types of networks that include end-to-end complex wide-area networks and sub-networks including local area networks, ad hoc networks, sensor networks, vehicular networks, and optical networks. It also includes research on heterogeneous networks that are hybrids of two or more types of sub-networks.

(Website: http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=12765&org=CNS)

- **FIND:** FIND (Future Internet Design) is a major new long-term initiative of the NSF NeTS research program. FIND invites the research community to consider what the requirements should be for a global network of 15 years from now, and how we could build such a network if we are not constrained by the current Internet -- if we could design it from scratch. FIND solicits research across the broad area of network architecture, principles, and mechanism design, aimed at answering these questions. The philosophy of the program is to help conceive the future by momentarily letting go of the present - freeing our collective minds from the constraints of the current state of networking. The intellectual scope of the FIND program is wide. FIND research might address questions such as:
    - How can we design a network that is fundamentally more secure and available than today's Internet? How would we conceive the security problem if we could start from scratch?
    - How might such functions as information dissemination, location management or identity management best fit into a new network architecture?
    - What will be the long-term impact of new technologies such as advanced wireless and optics?
    - How will economics and technology interact to shape the overall design of a future network?
    - How do we design a network that preserves a free and open society?

(Website: http://www.nets-find.net/)

5 EXPECTED NETWORKED MEDIA OF THE FUTURE

The Networked Media of the Future is built upon three axes that will be available in the future, i.e. true broadband, personalised media, and distributed control. If these axes are fully develop the **Networked Media of the Future will provide new models of rich media interactions and cooperation based on enhance AV content at the community level with distributed control infrastructure and community intelligence allowing pervasive personalised media services without the need of central control**.

This vision of pervasive personalised media services without a central control requires research in several domains. In terms of *content and services*, there is a shift towards enabling non-professional production of content, creating services that reflect to societal needs, and emergence of multi-party, multi-play services. In terms of *access*, people are expected to retrieve their multimedia assets and play them on their devices while being mobile. In terms of *collaboration*, technologies need to leverage the increasing computing power, data capacity and connectivity of surrounding devices, and exploit a distributed fashion, as well as to enhance *virtual reality and conferencing* technologies to enable a more natural and effective team collaboration, and also *empower virtual communities* to create user-generated content and services.

Consequently, research need to focus on creating a sophisticated environment, where all of the human senses are engaged (but not overwhelmed) and where communication via technological means preserves (and maybe augment) the richness and subtle characteristics of human-to-human communications.

This section is structured around three axes espress in terams of questions by the three FP6 Networks of Excellence.

5.1 The challenge of true broadband

This challenge may be summarized as **"The challenge of true broadband",** as a new situation in which Gigabits per second (or even higher) connectivity will be supported to not only our homes, but also wherever we go/travel and whenever we would like soon. The real question here is indeed: *"what will we be able to do with this very high-speed connection, which we cannot do now?"* To answer such demanding questions, it is firstly necessary to identify the applications that can benefit from such high speeds as well as their associated research challenges. **Figure 9** illustrates our envision on what the real research challenges will encompass to provide and then exploit true broadband at anywhere with an ultimate potential for supporting true multimedia experiences to its users in the future, as addressed by the VISNET II NoE question. This question has been analyzed in the Networked Media Task Force from two different points of view: the application environments and the enablers needed for these environments, as depicted in the figure below. Both points of view are detailed in the following two subsections.

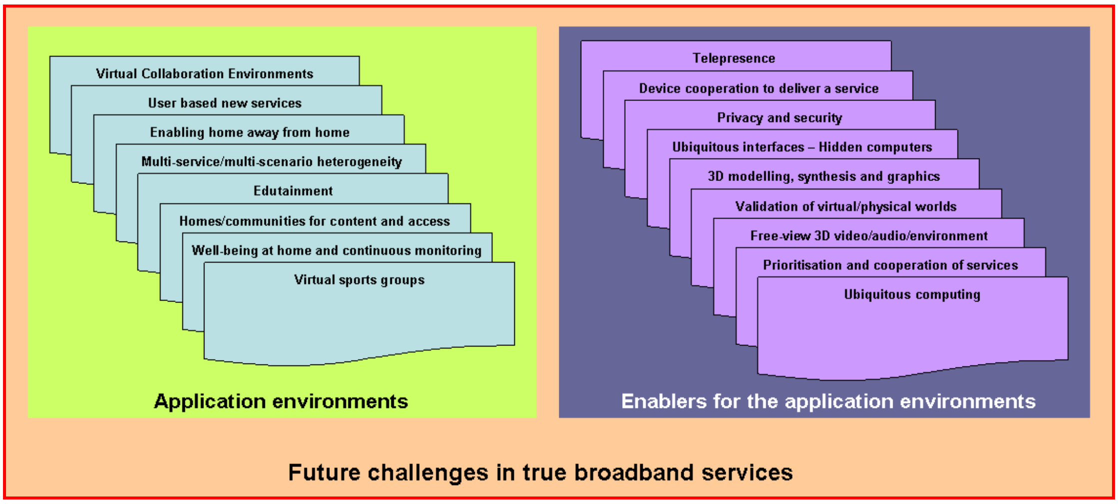


Figure 9: The challenges to address for realising true broadband

5.1.1 Environments

5.1.1.1 Virtual Collaboration Environments

A future world where people can collaborate together effectively in teams, without needing to be physically co-located will be essential to meeting a range of work-place related tasks, managing transport congestion and responding to environmental pollution. Hence the goal for Virtual Collaboration is to enable people, who are geographically remote from each other, to collaborate where the standard of collaboration among the parties is of a quality, richness and subtlety comparable to what would obtain if all the parties were co-located in the same room.

To achieve this the nature of the social exchanges that make up a typical collaboration will need to be studied and analysed where the analysis of an encounter should take into account not just the participants and the content of what was exchanged but also the location. The interaction of the sense of place with the more subtle protocols of communication leads to the view that a model of the physical space in which people interact should be produced and against this background various types of communication can be modelled.

The typical organisation of a traditional office building provides a sensible basic metaphor for the electronic implementation of the sense of physical space hence each virtual collaboration space should have a definite boundary. The preservation of geographical integrity should be a basic principle and, consequently, users of the space should be allowed to electronically navigate the space at a typical walking pace (i.e., if it takes 10 seconds to traverse the physical space it should also take 10 seconds to traverse the virtual space).

In such a space, no implementation of vision or hearing should be allowed that would not correspond to what would have been possible and reasonable in the real world. Thus, what a person could "see" or "hear" electronically should not exceed what he/she could have been seen or heard in the corresponding physical situation. Socially unacceptable behaviour such as eavesdropping should not be any more tolerated because the transgressor had his/her ear to an electronic rather than a physical keyhole. People have a strong natural sense of place and the degree of privacy and the type of behaviour that would be appropriate to different places.

There is little doubt that the Virtual Collaboration will gradually work its way into the public and political consciousness as the natural and environmentally virtuous way to reduce society's dependence on the physical transport technologies. It would however be foolish and short sighted to fail to recognise the reality, sophistication, scope and potential of the various enabling technologies, and that they will inexorably converge to provide ever more natural communication.

5.1.1.2 User based new services

Two key trends in Internet usage patterns have come to the fore in recent years encapsulated under the terms User Generated Content (UGC) and Web 2.0. These phrases are in relatively recent circulation (Web 2.0 was coined by an O'Reilly conference in 2005) and their precise meaning and impact are still evolving. However, they both represent a clear change in the way the Internet is used – it has evolved from being a read-only medium to being a read-write medium. Users now expect to contribute to an Internet experience and benefit from other contributions, as well as access more traditional content.

Currently UGC is typified by personal blogs, community news sites such as Slashdot, photo sharing sites such as Flickr and community knowledge sites such as Wikipedia and new media sites such as YouTube. The rise of YouTube in particular highlights the desire by users to be both consumers and creators. YouTube also highlights that new media is a powerful force and that news journalism cannot rely on the written word to reach the new, and even older, generations. MySpace may typically be thought of as a friend's network but it shows the power of music to bring people together. Flickr demonstrates the appeal of static images while Wikipedia and blogs tell us that the written word is still important when coupled with subjects that people are passionate about and in a format and process that is anathema to old world journalism. A key aspect of UGC is the aggregation of attention. Through the minute actions of millions of users YouTube finds and features the most popular videos. This feeds a loop that drives popular videos even higher. Where before a single or closed group of arbiters decided what users hear, see and read now users are rejecting this and choosing for them. Users are becoming both consumers and participators Web 2.0 fits into this by being the technological and conceptual framework in which UGC flourishes. Web 2.0 espouses open standards but even more than that, open systems. It goes beyond the hardware and software by encouraging businesses and people to open their content up to be distributed and used in ways not originally envisioned. Content is treated on its own merits rather than through which medium or publisher it came from. Users are given the power to select from disparate sources and aggregate into one view what they want to see. Systems are encouraged to open up so that users can work them together in new ways; for instance combining mapping services (such as Google Maps) with data sources (e.g., crime figures, property prices, traffic data etc) to generate a new type of service.

5.1.1.3 Enabling home away from home

These trends imply that people are able to get lots of information from outside their home, actually from all over the world. As the network bandwidth has been ever growing, both the quality and quantity of information and services which users can access also have been improved from simple texts and pictures to voice, video, and 3D contents. People can also use these services even outside the home with the help of various mobile devices such as mobile phones, PDA, multimedia players, portable game devices, etc. However, the challenging problem is that the required bandwidth is always over the provided bandwidth in current network environments, which incurs delayed, less scalable, and poor quality of services. That is the reason why researchers have devised many approaches that could reduce the bandwidth requirement without degrading the quality of service in various fields such as multimedia streaming, P2P file sharing, scalable transport protocols, and so on.

If we can get high speed connection with our mobile devices which provide us with multimedia services, we still have challenging issues in two aspects: quality of service and timeliness. The more bandwidth we have, the more quality of service we get. Video conferencing with friends could be done via more realistic and high quality video of the counterpart. It would be possible to serve real-time movie streaming with the DVD quality even while users are travelling. As the quality of multimedia grows, users and service providers also requires more bandwidth, which implies that the research challenges could be still how to support multimedia services with the highest quality, because even gigabit wired/wireless network could not satisfy the quality requirement in future.

5.1.1.4 Multi-service/multi-scenario heterogeneity

The future digital world will be characterized by the emergence of a social infrastructure where access to knowledge resources and services is no longer location or time dependant. Today there is a strong coupling between services and communications infrastructure used to deliver these services. Users are interested in using services and accessing information, not in accessing nodes that hosts information or provide services.

In the future, the emphasis will be on services which reflect societal needs. Increasingly such services will be produced and customized by non technologists who will have the ability to create new services based on the integration of services and content from a multitude of services.

As system, network and service complexity grows, value chain branches will take new shapes. New roles will be created, merged and or split with existing ones in continuously new configurations. Users' behavioural patterns, movements, locations, choices and actions contribute to service creation and personalisation. End to end services will be dynamically generated, integrated with existing services and delivered across organisation and network boundaries. Every link in this value chain relates to an entity that trades with other entities in order to meet the consumer service needs and quality of service. Service providers will need to develop new software applications that adequately support service lifecycle management, based on context awareness, personalisation, and intelligent dynamic service composition. Service Providers establishing multiple relationships with software providers and consumers in a highly complex relationship where consumers of services may also be consumers of services creating new forms of co-development. Software providers and software developers, in their turn, address other players that develop and provide frameworks, methodologies, standards in order to perform, generate and manage their product-lines and services. These knowledge, product and service exchanges may adopt different models: purchases, medium- or long-term agreements, business rules, open source and community paradigms.

Other issues that need to be addressed in such an environment include the issues of service discovery, service indexing, service update and evolution. Other challenges will include the development of distributed multimedia knowledge-based management systems, multimedia data-mining, indexing and context driven adoption.

5.1.1.5 Edutainment

Edutainment is one product of a major shift - tied to changes in the economy - that is occurring in how we view leisure time in today's modern societies. Nowadays, more people work with their brains than their bodies. People are using their scarce, but more highly valued, leisure time differently, and they have an entirely new attitude about leisure. They see leisure time as an opportunity to improve themselves and their children and do worthwhile things, rather than as purposeless relaxation and entertainment. With the emergence of a ubiquitous wireless communications environment mobile television and particularly interactive IPTV is seen as one of the key platforms for the development of Edutainment products and services. While the market has mainly been focused towards children learning through play future systems will be provided to all ages providing access to a broad range of services from music through travel to more traditional education content. Issues which need to be considered include personalization, localization and pervasive computing schemas in additional to the pedagogical issues. The inclusion of Virtual Reality into the design of Edutainment products will provide opportunity for the user to slip in to a role such as a surgeon, pilot or explorer and fully experience the role. The creation of sophisticated 3D virtual reality environments which can mingle and engage with virtual tutors or visit some of the great museums or even revisit a time in history will be possible in near future. A key challenge is to address issues such as users' senses and emotions as well as the ability of users to personalize and control such environments.

5.1.1.6 Homes/communities for content and access

With the help of high speed network, it would realize the distributed large content creation, exchange, and management among communities. A user could use mobile devices to access and manage his/her data which is distributed in computers of his/her home and office, which will be one of useful applications. It can be realized because the enough network bandwidth could compensate the limited storage of mobile devices. This brings an issue of how to synchronize distributed data in real-time when it is managed in a remote way as data stored in one place could be managed (updated, saved, deleted, etc.) in another place by possibly multiple users.

Another promising application could be remote control and processing of content through which users with mobile devices leverage remote high capacity machines and receive the result of processing and control. This is similar to VNC and remote desktop. However, it will be so improved that a mobile user can play 3D games that can be executed in his/her home PC even if the mobile device does not have 3D rendering capability. The reason is that it is possible to transfer the whole screen image from his/her home to the mobile device in real-time. It gives a user an impression that he/she plays the game executed in mobile devices.

In addition, we could combine the use of distributed large data and remote processing techniques as mentioned above in order to solve heterogeneity problems in a way that a mobile device only presents contents which is distributed and stored in several different places and it uses a program executed in another device to present the content. Research challenges will be how to allocate and distribute content in an efficient way and share nearby processing power in different computing devices when there are concurrent resource requests.

5.1.1.7 Well-being at home and continuous monitoring

Tele-care services developed on from the Virtual Collaboration Environment have a large potential for health improvements by effectively and efficiently developing integrated services for health and social care in the home and community. These can include care of the elderly, including overcoming loneliness, people with long-term chronic health conditions, disabled and disadvantaged groups. Event-based home monitoring sensors and monitors could feature whenever appropriate and remote diagnosis and care services can be enhanced using high definition video-conferencing. Such services would enable home-based training, treatment and follow-up for the elderly, and enabling the elderly to live independently at home for longer.

Home based e-health services can give improvements in areas such as stabilisation of chronic disease, reduction of delayed injuries and reduced hospitalisation with improved quality of life, knowledge and safety. Furthermore, the current health system will not be able to serve the increasing elderly population by hospitalisation; the home has to be used as an arena for e-health services in order to serve us all in the future. The new procedures, services and technologies so developed will enable health care personnel at hospitals and care institutions to seamlessly supervise, follow-up, treat, and monitor elderly people in their own homes. Further research and development through universities and health care providers would be enabled with these initiatives being built on the expertise extant in the community. However, great care must be taken with privacy and security of the information relating to these services as they are being designed and implemented for some of the most vulnerable in our society.

5.1.1.8 Virtual sports groups

Enormously improved network bandwidth enables both sports and game areas to benefit from it for efficient player management and real-time multiplayer sports game. For the player management in real sports, with various embedded sensors, trainers and coach staffs could keep track of players' physical condition and record their whole play in order to analyze and improve their abilities. With this large amount of useful sensed data, players, trainers and coaches could improve play skills, strategies, training plan, and so on as they can know how a player's current physical condition is and how it is changing in real-time based on more accurate quantitative information. One of challenging issues is that it requires conveniently wearable and more powerful sensors as they should be simply worn by players with long battery life as well as affordable processing/networking support inside them.

Another promising field is online sports games. There have been already lots of massively multiplayer online role playing games (MMORPGs) where thousands of simultaneously connected players play together. On the other hand, for sports games, much less number of players (from two to six players) is allowed to play with each other because the sports game is different from MMORPG such that it needs more timely interaction with relatively faster playing pace among players, which requires enough bandwidth for real-time message transmission. Of course it is not solved only by providing high-speed networks as recent games also require lots of CPU processing power for rendering 3D scenes and performing a large amount of game simulation and calculation. Therefore, it is also important for the bandwidth to compensate for the lack of CPU capacity.

Finally, techniques to assess the confidence in the privacy and security put into place, particularly for dynamic and complex infrastructures, need to be established.

5.1.2 Enablers

5.1.2.1 Telepresence

Although the technologies that underpin telepresence have been available for some time, the applications that would fully justify the usage of telepresence (such as medical applications, remote operator applications, advanced teleconferencing applications etc.) still do not utilise such technologies fully. Although MPEG-7 provides a basis on which to define the content and the organisation of content, it is far from satisfying the needs of highly specialist applications as detailed above. Considering the difficulties involved in tailoring a new set of tools for enabling applications in each of the mentioned areas, it is necessary to devise a unified framework that is independent of technology. Namely, guidelines should be generated which will quantify the requirements (such as sound source localisation acuity, the weight of the haptic interface in the application, the method, resolution, and acuity of 3D video reproduction etc.) In addition, strategies must be developed to allow broadcast telepresence to be experienced by multiple users.

Research efforts on Telepresence have been strengthened worldwide recently, covering the whole media processing chain from capture to display. The goal is to extend visual sensation to the third dimension in order to achieve the impression of presence in communication with remote sites. This technological challenge is usually based on 3D vision and improved by realistic simulation of the other senses. 3D vision simulation is inspired by the vision system of biological creatures. It is based on the two eyes and the brain, where the eyes capture aspects of the 3D physical information and the brain recovers the distance or third dimension. In 3D vision, telepresence aims at simulating this system using two or more cameras for capturing 2D projections – images – of the physical 3D world-scene. A computer takes the place of the brain in the computational modelling, processing, and interpretation of the 2D projections. This is the task of stereo vision: the recovery of the original physical 3D information – depth – in a manner similar to the way biological creatures perceive depth. Once the depth information is available, it is possible to represent the 3D information in a flexible way. These representations have critical applications in medicine for 3D visualizations of internal organs for inspection and diagnosis, 3D teleconferencing and many others.

Unfortunately, 3D vision has not found a killer application yet. It is rather well-established in niche markets, including professional applications as the mentioned scientific visualization in medicine and entertainment (IMAX® cinemas and 3D gaming). A reason for the low penetration of existing 3D vision and telepresence technology can be traced back to several unsolved problems that render sub-optimal the visual performance of available commercial products. However, in recent years research efforts have been strengthened worldwide to remove technological obstacles impairing a wider penetration of 3D video and telepresence applications.

Important technological breakthroughs are needed in all components of the processing chain, from acquisition and signal processing, over 3D scene representation, coding and transmission, to rendering and 3D display. Very likely various 3D video systems, components, applications and services will enter the market in the near future.

The following outlines important areas of work related to telepresence and 3D vision that need to be addressed in a concerted effort at European level.

STEREOPSIS AND STEREO VISION

Stereopsis is concerned with the 3D viewing of objects. This problem is studied in different branches of knowledge, including art, architecture, philosophy, physics, psychology, neurology, amongst others. The main target in early vision was to develop constraints to make the problem well defined and solvable. Stereo vision is an interdisciplinary science. The main problem of stereo vision is to understand the scene structure and its 3D properties from a set of images taken from different viewpoints. The stereo vision problem – from a computational perspective consists of two main problems: estimating the parameters that encode the scene geometry – known as the calibration problem – and identifying the position of pairs of image points such that each pair is the projection of the same 3-D scene surface point on both images – known as the correspondence problem or disparity estimation. For estimating the parameters that encode the scene geometry, it is usually assumed that a set of corresponding points is available. On the other hand, in order to reduce the search space and obtain a more reliable estimation of corresponding points, it is required that the images are calibrated. Once the geometry of the stereo system is known and corresponding point estimations are available, this information is used to compute the 3D location of the points from their projections (reconstruction).

Fundamental research is still required to solve these two problems with a high degree of accuracy and reliability. A critical aspect is the complexity of existing technology. Much more efficient algorithms need to be developed to cope with the large amounts of data that need processing in real time for low-access latency in telepresence systems.

IMAGE SYNTHESIS

The last three decades have seen rapid developments in 3D computer graphics. The process of converting a model of a 3D scene into a desired image is a well understood task and there exist several well established algorithms. Image synthesis applications have to deal with the complexity of the real world by combining the available reference images with the given scene geometry knowledge. In 3D computer graphics, virtual images of a 3D scene are created using geometry-based rendering. The 3D scene is represented by explicitly constructed 3D model and virtual images are produced by changing the model or virtual image position and rendering. On the other hand, image synthesis concerns the generation of virtual images of a 3D scene from its 2D representation, without using a 3D model of the scene. It requires an accurate level of scene geometry information. The scene geometry information can be obtained by disparity estimation. However, the level of realism of a virtual image depends on the quality of disparity estimations. Some real-time applications require the use of sophisticated purpose-built hardware accelerators due to the computation cost for obtaining high-accuracy disparity estimations.

The creation of virtual images requires high-accuracy disparity estimations to produce high-quality virtual images. Disparity estimation is the solution of an inverse problem. There is a lack of information about the 3-D scene structure that makes this problem an ill-posed problem in the original sense of Hadamard. It means that the solution is highly sensitive to perturbations. However, stereo vision systems embody the assumption that the scene was captured with precisely calibrated cameras and the ideal disparities are available for any point in the 3D scene. The problem is that the ideal disparities cannot be estimated and consequently small distortions in disparity vectors lead to significant distortions in 3D scene coordinate estimations. Moreover, distortions in the disparity vectors produce artefacts in a virtual image. As a consequence more research and implementations in the area of image synthesis is still required to obtain reasonable results.

CODING OF MULTI-VIEW VIDEO, DEPTH AND ASSOCIATED DATA

Many telepresence systems are based on stereo and multi-view analysis. Depending on the degree of common content, shared by a subset of the cameras, a coding gain can be achieved in comparison to single-view coding. In multi-view coding, correlations between adjacent cameras are exploited in addition to temporal correlations within each sequence. Thus, multi-view coding adds another compression dimension on top of single-view coding. Stereo view is the most important special case of multi-view analysis. In stereo the images are in general very similar, which makes them well suited for compression, e.g. with one image predicting the other.

An alternative to classical stereo video coding is to transmit a video signal and a per sample depth map. From the video and depth information, a stereo pair can be rendered at the decoder. This extends the functionality since it enables head motion parallax viewing if the user’s head motion is tracked.

Research on coding of stereo video, multi-view video and associated depth or disparity data has already reached a good level of maturity. However, compared to coding of other types of media data there is still a lot of room for improvement of algorithms. This includes for instance optimization of depth or disparity coding by dedicated algorithms. In this context integrative research at European level would be very useful to make real in-roads in the development of a more efficient new generation of coding algorithms for telepresence applications.

5.1.2.2 Device cooperation to deliver a service

The wide availability of devices with processing, memory and networking capabilities has led to the development of novel inter-networking and distributed processing strategies, which take advantage of the cumulative power of these clusters. Two of the most successful examples of these approaches are peer-to-peer (P2P) networking and grid computing. The former consists of a decentralised networking infrastructure where nodes are connected in an ad-hoc manner. It can be used to share files (data and multimedia), multimedia streaming and voice-over-IP communications. The service delivery relies mainly on the bandwidth and computing power of participating nodes. With the ubiquity of wireless networking, we are witnessing P2P with portable devices, such as game consoles and mobile phones.

Grid computing is another example of distributed architecture where the main purpose is to offload the main processing tasks to more powerful machines or to access large amount of data. This is particularly useful for media processing and distribution. These architectures can be scaled down to home size (e.g. streaming video from PC to TV or portable device), and up to buildings and public spaces.

Furthermore, node devices closer to the end user may collaborate to decide on which is the best way to deliver the service by e.g., balancing quality and mobility. In another scenario, multiple devices may support the different modalities of the service; for instance, use the computer to order the movie, then stream video to the TV, and audio to the surround system while using mobile device as remote control.

5.1.2.3 Privacy and security

Many of the proposed environments, including virtual collaboration, require the secure distribution of multimedia data or resources amongst a group of users. These users may be on many different networks, using many different devices, belonging to many different organisations (e.g. for multi-company collaboration) with different security solutions. The groups may also need to form and adapt dynamically while still maintaining security. Applying security in such heterogeneous environments requires new solutions. Security management solutions are also needed, with the right level of centralised versus decentralised management depending on the scenario and being appropriate for the technical skills of the users involved (which could be very low). This includes the need for multi-domain authentication and authorisation models, building on and enhancing existing work (such as Federated Identity Management and grid computing security) for the new multimedia environments being considered.

Users should be allowed to control their identity, personal information etc. especially in a world where context related privacy, such as user location, will become increasingly more important for more advanced multimedia applications (e.g. virtual presence or telepresence). Last but not least, the issues of end-users distributing their own content and affecting the privacy of others (e.g. using a camera phone to record images of a third party, which are then posted on the Internet) must be addressed. In the situation where end-users are becoming producers of content makes it harder to assess if content is trustworthy.

Techniques to set-up and manage security dynamically across heterogeneous networks, in a way compatible with content adaptation must be established. This should also include access control management for dynamic groups.

Finally, techniques to assess the confidence in the privacy and security put into place, particularly for dynamic and complex infrastructures, need to be established.

5.1.2.4 Ubiquitous interfaces – Hidden computers

Since Mark Weiser coined the term “ubiquitous computing”, enabling technologies have been incorporated in the design of hardware/software, user interface and networking. This has resulted in the creation of smart devices, embedded sensors and always-on networks, etc. In this vision, objects exhibit computing capabilities and collaborate with each other to deliver a service to the user. It is expected that people would interact with these “invisible” devices more naturally then with personal computing.

For the realisation of this vision, the different facets of ubiquitous computing are being investigated. In the computing world, there is move towards distributed processing (e.g. grid computing), distributed software architectures (e.g. web services, P2P) and component-oriented programming paradigms. Alongside processing, the connectivity is an essential component for the system. In the networking world, we witness a convergence of the next generation networks, wired and wireless as well as fixed and wireless, with the aim to provide always-on connectivity to everybody.

The distributed nature of the interaction and their embodiment in everyday objects means that conventional input modalities (keyboard, mouse, etc.) are obsolete. Some progress has been made in the creation of new interaction techniques (scroll, stylus, voice, etc.) for portable devices (PDAs, mobile phones etc.), resulting in incorrect interpretation of the input commands. Consequently, advancements have been made in the creation of interactive surfaces, where elements from the user’s environment can serve as input/output devices; ambient interaction, where sound, colour and light react to external stimuli; and tangible interaction, where the manipulation of physical objects is reflected in the interaction with data. These developments became possible thanks to the advancements in gesture recognition and sensor technologies. Besides the tangible and haptic interactions which rely on physical contact, audio is potentially more suitable for ubiquitous interaction because it does not require the direct attention of the user and it is closer to human-to-human interaction. Examples include voice recognition, speech synthesis, and auditory notifications.

This post-PC shift raises numerous challenges for product designers and user experience evaluators. They result from the interplay of the complex technologies, the physical design of the interface, the information flow between the user and device, and the context in which it takes place. The new way of interaction bears little resemblance to the traditional approach as the systems take a more pro-active role in anticipating and interpreting the user actions. The implementation of such systems requires a high level of intelligence whereby the system can understand the commands and infer their relationships to the current context. Other practical issues relate to how the system should provide feedback about the successful (or failed) execution of the commands, and what is the best way to convey a command to a distributed system.

These developments denote the increasing importance of the human user. In fact, the human body, as a whole, is becoming a medium for interaction. By capturing and interpreting these signals (physical, movement, physiological, etc.), we will be able to achieve a very natural interaction with the ubiquitous system. This suggests the use of different sensors on the body and even inside it. The introduction of ubiquitous computing will penetrate into the close vicinity of our body, such as our clothes, rooms, but it will also be extended to buildings, public spaces and ultimately everywhere [22].

5.1.2.5 3D modelling, synthesis and graphics

As the commercial success of Second Life demonstrates, 3D worlds now enter a new era. Authoring personal 3D information becomes a usual matter and is spreading out to non specialists.

One main limitation of 3D modelling and the displaying of 3D content is that it consumes a lot of processing power. With the boom of graphics chips, the technical possibilities are growing in a lot of different directions. Thus, running those applications in an extended home environment is problematic due to the large amount of potential platforms the applications have to function on. Given gigabits per second connectivity, authoring tools could compute demanding operations on the server side thus allowing the use of less capable machines at the client side. In this light, displaying complex, physically based simulated scenes could be achieved on any kind of hardware with simple display capabilities. Indeed, with a bandwidth big enough, the frame buffer could be computed on the server side and sent to the device at a given frame rate.

Projects such as Folding@home and Seti@home demonstrate the potentially available computation power through network based CPU sharing. But for real time purposes, its usage is limited due to restricted bandwidth.

5.1.2.6 Validation of virtual/physical worlds

Over the past fifteen years, virtual worlds have become more and more realistic. Video games, 3D animated movies or virtual communities have made interaction with 3D environments increase in popularity. The technique is now at a level where computer generated image can be mistaken for real. But to reach such a level still requires the skills of a gifted designer.

Automatic approaches to model the physical world exist but require too much computing power to run on non-dedicated hardware in real-time. Thus an approach made possible by Gigabits per second connectivity is to make the computation on the server side and transfer the outcome.

One further research challenge the research community has to overcome is with respect to real-time, realistic, physically based simulations. E.g. cloth or water simulations allow us to reach a level of visual accuracy comparable to the reality, but still some simulation domains, such as e.g. hair or face simulations, have to improve to reach a level of believability that makes it indistinguishable from real scene footage.

Furthermore, the accuracy and realism of virtual environments need to be verified in two respective ways. The physical aspects of virtual environments are easier to validate and reflect the accuracy of the model in comparison with real-life conditions. The perceived realism of such environments is harder to quantify and validate as there are many unknowns, particularly regarding the interactions between the different modalities. Namely, the precedence of different sensory cues and their interactions needs to be quantified. In addition, metrics and models for the perception of virtual worlds need to be developed.

5.1.2.7 The challenge of free-view 3D audio-visual environments

A variety of 3D audio rendering methods exist and the If free-view 3D audio-visual information and the environment are provided accurately, they will then be significant for enhancing the user experience of the multimedia presentation and/or telepresence systems without doubt. Users of such systems would certainly like to be fully immersed into the 3D scenes both visually and in auditory terms while interacting with the presented content as well as each other. For instance, while watching a sports event, one would like to see or hear from an angle of his/her choice. This feature evidently requires an interaction with the presented scene. Interactions particularly within multi-user systems, such as virtual collaborative environments, may be multi-modal to provide services effectively. Free-view provision of the services enable those users to view, hear and/or interact with the accessed/distributed content in line with their preferences/needs using mixed and/or virtual realities, which in return will enhance the quality of user experience.

There exist a variety of 3D audio rendering methods and the area can be considered as very mature regarding the availability of different reproduction methods and different setups for different occasions. In order for a 3D audio system to provide free-view, it should essentially provide in a large area for one or more listeners:

(1) A reasonable accuracy for localisation at the ear level (i.e. the azimuth cues),

(2) The height of the sound source (i.e. the elevation cues),

(3) The distance of the sound source.

More often than not, existing methods fail to provide either one or more of the mentioned cues, and mostly only for a single listener or for a small optimal listening area. The audio research that needs to be carried out should provide new rendering methods, novel setups, or different combinations of existing setups to overcome the existing limitations. The resulting techniques may either be physically accurate as in WFS or ambisonics, or provide perceptually acceptable results as in conventional multichannel systems with the condition that the listening area is enlarged.

Similarly, since 3D free-view video representation constitutes an important part of enhanced interaction between the user and the 3D world, advances in 3D free-view video should be considered. The users are expected to watch the scene from any viewing position while being able to see every detail in the scene which would only be seen if the camera was located at exactly where the user stays. Different approaches, some of which are still under research, exist in order to enable the synthesis of arbitrary view-point pictures to be put on displays. Applicability to any scene, regardless of its content and visual complexity, ability to generate high quality virtual pictures in a large span of angle with the minimum dependency on available camera locations and arrangement, fast rendering and fast switching between rendered views, should be the main concern of a novel 3D view rendering scheme. Reliable communications of scene depth information is essential for the generation of high quality and realistic representation of arbitrary view-point videos to be presented to the user. One of the current and so far the future scopes of multi-view video coding research is the realistic generation and reliable transmission of such auxiliary video data for the realisation of 3D free-view video. On the other hand, applicability of the system to multi-user scenarios is also important which makes research on advanced 3D video displays essential. Currently, 2D or multi-scopic displays using head motion tracking mechanisms exist which restrict the number of interacting users. Displays, showing different scenes from different viewing angles simultaneously, should thus be constructed.

5.1.2.8 Prioritisation and cooperation of services

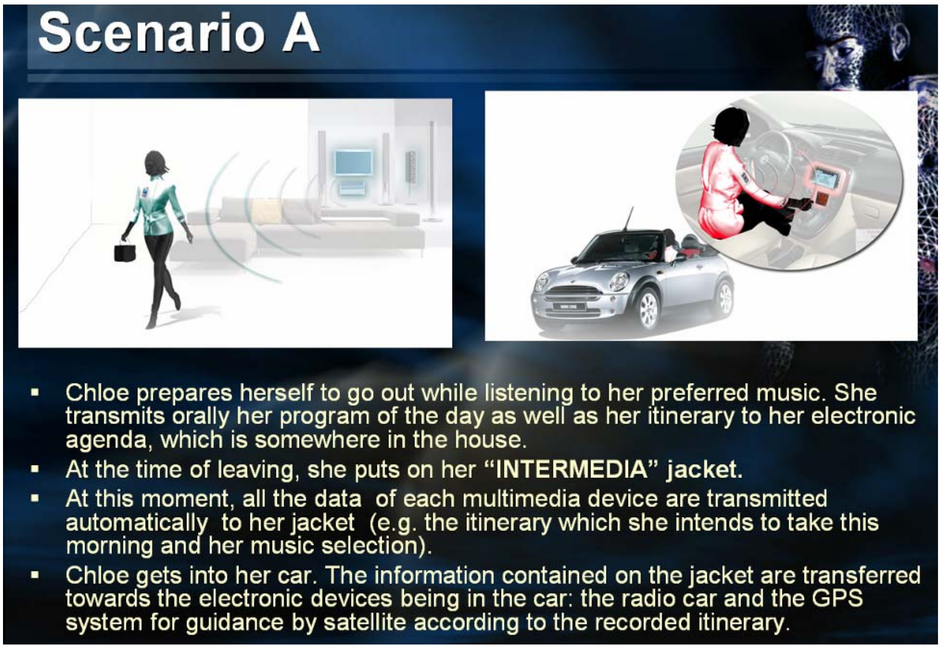
There is a growing trend in software architecture to build location- and platform-independent services (e.g. web services), which makes it possible to build complex services out of simple components. Thus, it will become easy to support scalable service delivery depending on the device capabilities.

The separation of services based on component frameworks of services reduces the complexity of development and enables the creation of composite services. Besides, the availability of open APIs to these services and appropriate software frameworks will encourage novice users to create and personalise these services. This best current illustration of the power of user generated services and creativity, when access and APIs are made available, is exemplified by numerous applications based Web technologies (also called Web 2.0).

Certainly, the ability to compose a multi-play service out of other service makes it easier to select the suitable ones in response to the current context or user preferences. However, “Quadruple play” (as the wireless extension to triple play), to deliver video, Internet access, and voice telephone service, will require a mechanism for adaptation and prioritisation of the services, especially when simulation performance is less predictable, devices are heterogeneous and simulation overall running computational cost is higher.

5.2 The challenge of personalised media

"The challenge of personalised media" is a new situation in networked media. The real question here is indeed: *how recent advances on relative to networked media disciplines can influence them?* To answer such a demanding question, it is firstly necessary to elaborate on the two most prominent key research directions as identified by the multidisciplinary NM-TF group: 1. Real-time and interactivity and 2. User domain and inclusion. **Figure 10** illustrates two scenarios on what the real research challenges of personalised media will encompass, as addressed by the INTERMEDIA NoE question. This question has been analyzed in the Networked Media Task Force from the above two different points of view and both points are detailed in the following two subsections.



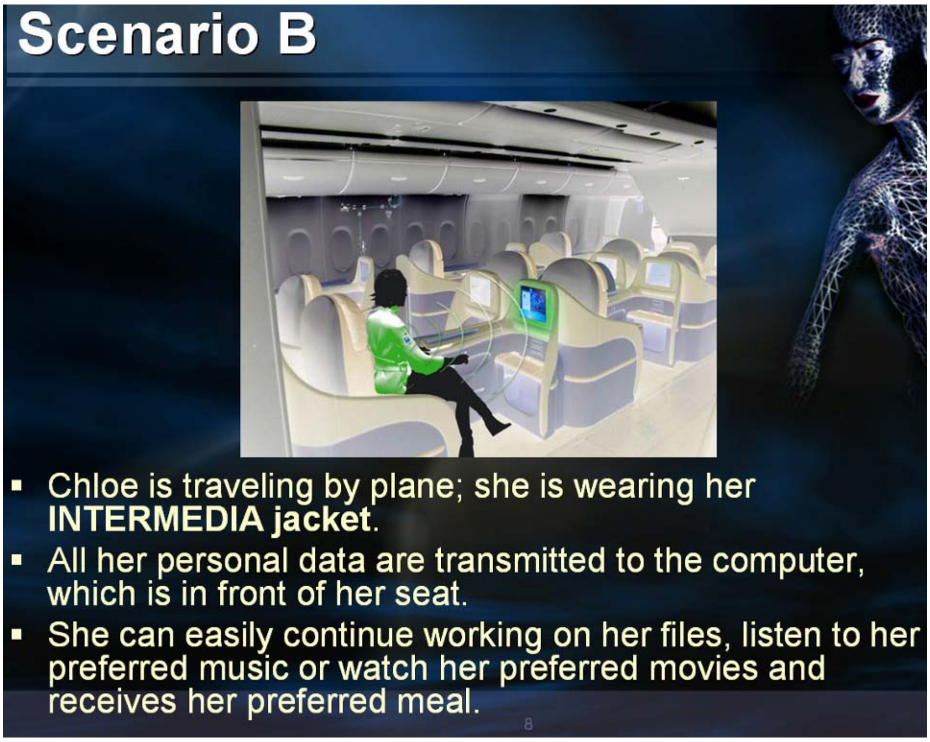


Figure 10: INTERMEDIA scenario of personalised media

5.2.1 Real-time and Interactivity

5.2.1.1 Compensate memory and computer power with bandwidth

High bandwidth allows sending more data and thus could be used to overcome clients' limitations in memory and computer power by feeding it dynamically with data. For memory limitations for instance, we could imagine that the local memory of a client is delocalized on servers and retrieve quickly through the network assuming high bandwidths. Let us take an example: the virtual tennis match. Two distant real players are playing tennis on a virtual court via their avatars and in a distant place (one player is in London and the other one in Geneva). In today's conditions, we need to keep equal capacity of all the information on each computer. What is transmitted is only the events update data. With more bandwidth, we could expect that only one site has a high-end powerful computer and that the other site could have limited computing capacity. It means that the limited capacity is balanced by the computer that calculates most of the simulation and a lot more real simulation data is thus exchanged through bandwidth and via the network.

Another crucial issue is on the "events data": i.e. to define a sufficiently expressive and powerful metadata associated to events in order to describe and communicate in real-time also the subtleties that are involved in communication. The communication in the virtual tennis scenario, for example, might include information, intentions, and expressions aiming at confusing the partner during the game, e.g. subtle moving and attitudes (e.g., aggressiveness). In other words, borrowed from theatre, it is important to exploit the "scenic presence" of the human by appropriate platforms considering the processing of appropriate arrays of cues on the non-verbal, physical, expressiveness etc. in the communication process, in order to achieve quality and the level of experience closest to the effectiveness of reality (but without necessarily trying to replicate it).

5.2.1.2 Interactivity with real-time

Interactivity requires a strong and practical back-channel from the client to the server and a well-defined loop of communications between them. For example, today we could have a TV real-time show via the Internet. This show is executed in real-time, which means that everybody in any place sees it as it is being broadcast 'live'. What interactivity could bring is the fact that anyone being on the Internet can follow the real-time show and take the floor and/or be filmed through a small web camera and be included in the show and watched in real-time by all others. Of course, it opens new avenues and questions as who takes the floor if many citizens like to, and maybe security and privacy as each citizen should login before in order to be identified in advance and so forth.

Another core issue in the TV scenario is on how to feed technology developments by strong and sound anthropological and social studies, theories from humanities, leading to novel models of interaction, novel "TV formats", thus defining the scenarios to implement new networked media platforms. The time scale of what is perceived as real-time in such new scenarios is a central issue: that is, novel real-time, interactive TV shows might emerge also by exploiting the constraints of such technologies. Let us consider the following example (borrowed from XVIII century Venetian music school of composers, e.g., Gabrieli's). They produced music masterpieces for double choir and orchestra, where the two choirs were located at a distance of several tens of meters, in baroque Venice churches. This implied a significant latency between musicians, between the two choirs playing together but in two opposite wings of the church. This was solved by innovative music compositional techniques by Gabrieli's exploiting this latency and reverberation constraints. This problem is analogous to the current Internet distributed collaborative performances, and might be inspiring for shaping the emerging interactive real-time TV shows mentioned above having latency and other typical features. This requires, for example, networked media providing reliable and controllable latency and scalable A/V features, influencing (and being influenced by) new TV shows, or composition techniques analogous to the ancient music example mentioned above. It is not likely that "TV shows", as they are conceived today, will be replicated or simply adapted in the future Internet real-time TV scenarios. A cross-fertilisation will be necessary in the development of networked media technologies from one hand, and innovative content from the other.

5.2.1.3 Real-time transmission of models: needs of a managing scheme

Managing schemes for real time transmission of models have to address the following issues:

A scheme has to allow efficient incorporation of:

- Large multi-modal control/interaction signals
- Specific background models
- Synthesized models from queries to large 3D/haptic/audiovisual databases.

The requirements for "real time" might vary in orders of magnitude (e.g. for haptic vs. audiovisual information). The managing schemes thus have to be much more general than current hierarchical and adaptive transmission schemes for large 3D models. Moreover, the managing schemes have to address the issue that there will be distributed sources (control signals of participating users, distributed 3D/haptic/audiovisual databases, background models) involving different systems and platforms, so that the question of new middleware arises as a subtopic.

In addition the managing schemes for real time transmission of models have to be suitable not only for secure transmission between different parties, but should incorporate notions similar to zero-knowledge protocols between the parties in order to ensure high privacy, as might be required by many applications (e.g. medical monitoring at home).

Similar to JPEG-2000 for instance, or other adaptive transmission schemes, we could devise mechanisms to transmit low to high level 3D/haptic/audiovisual contents. Scalable schemes for such media content would also be beneficial when having large database of content. Let us take an example in the clothing industry. Professionals in this branch like first to select a fabric by touching it. We would have people in Shanghai and in Geneva willing to buy fabrics. They will touch it over Haptics data glove and see it in 3D in real-time over the network. As the Haptics is sending 1 KHz over the network to the client instead of 25 Hz, we need new efficient transmission protocol which does not exist today, as we do not have enough bandwidth for this kind of an experience.

This example on experiencing "textures" may be extended to sound textures as multimodal, enactive experience embedding auditory, visual and haptics as a whole experience.

5.2.1.4 Transmission of large 3D/haptic/audiovisual databases

It is one of the central goals to create technical systems making it possible that different persons in different countries will be able to obtain the illusion that they interact (multi modally) in real time in a common 3D-VR-environment. This does not only include audio/visual interaction, but also real-time haptic interaction and appropriately sharing of haptic sensations. The state of the art with respect to sharing interactively a common audio/visual VR environment has reached a degree of moderate maturity. A fundamental problem in this context has always been raised: The improvement of methods to be used for audio coding and the efficient description and encoding of 3D models of humans or even of complete 3D scenes. All these efforts were necessary also for the purpose of transmitting the respective scenes to distant places (an example of such 3D warehouse is shown in **Figure 11**).

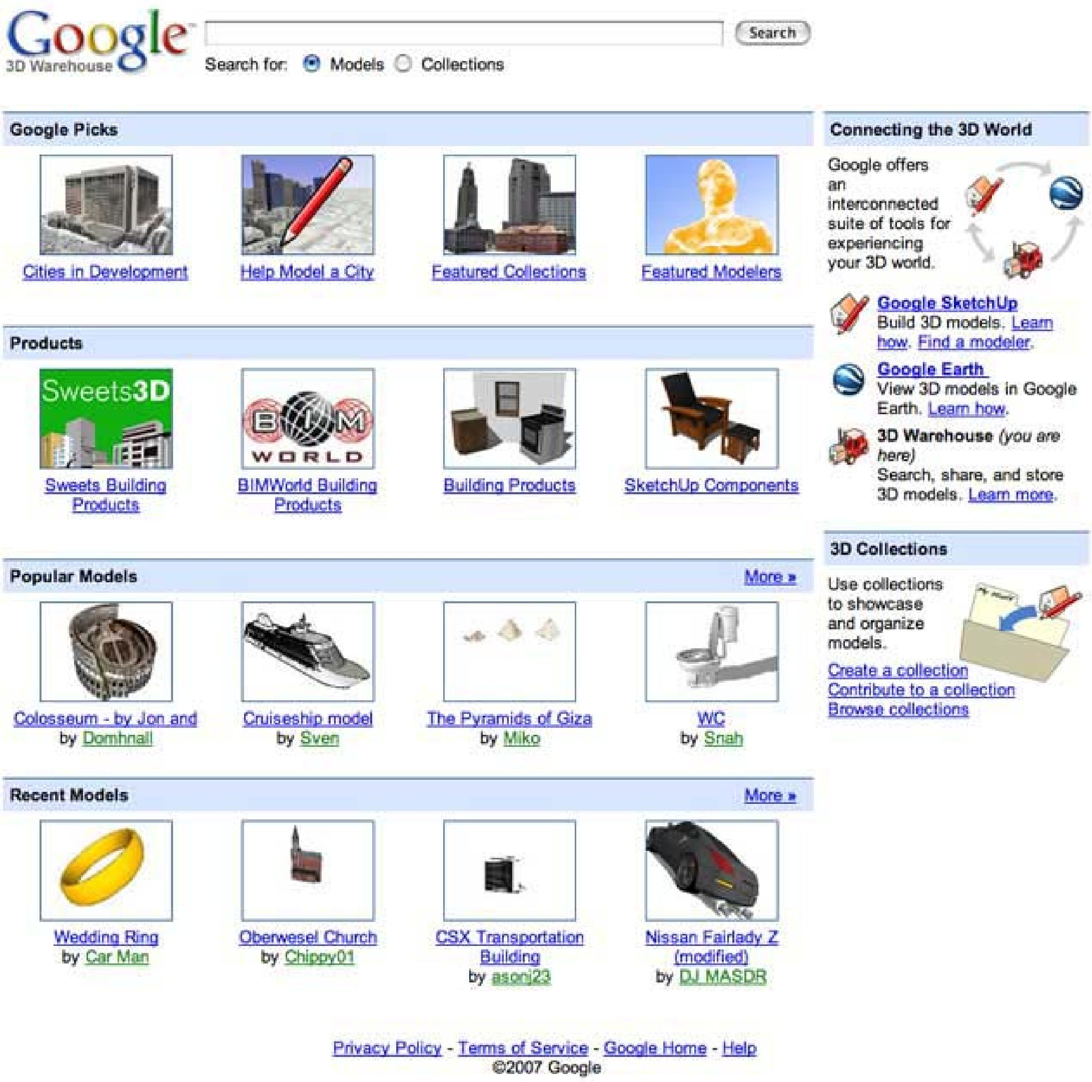


Figure 11: Google 3D Warehouse

Data Simplification based on Psychophysical Laws

In the context of the aforementioned (multimodal) interactive communication tasks, the psychophysical aspects of the respective tasks have always been crucial and will continue to be even more so in the future. The reason for this is that we want to construct multimedia systems transmitting signals evoking illusions or causing perceptions deemed as being realistic by the humans participating in the respective (possibly distributed) multimedia VR–System. This implies that it has been important to understand under what assumptions can 1) acoustic signals, 2) visual signals and in the future 3) haptic and tactile signals be considered as being equivalent with respect to human perception in some given environment or setting. Hence there exists a whole equivalence class of different acoustical signals creating for humans the same acoustical sound perception. The respective statements apply also for the perception of visual and haptic/tactile signals.

Due to the preceding observation, it has been a basic research problem to find the simplest signal in an equivalence class of signals creating the same perception. Here, simplicity means that it should be possible to create the respective signal with little effort and to describe the signal with as few data as possible. Achieving these goals for acoustic signals made use of psychophysical studies showing that in the perception of sound more energetic frequencies would dominate weaker neighbouring frequencies making them “unhearable”. Hence they could be ignored and omitted resulting in a much shorter description of an equivalent sound signal contributing to an efficient audio data compression. There are some well known yet important psychophysical insights that have been employed for visual data compression. Those insights include that due to the inertia of visual perception a sequence of not more than twenty images per second is sufficient to create the illusion of a continuous motion picture limiting the amount of data that must be transferred. The information describing colors can be compressed due to the psychophysical (tristimuli based) RGB-model (employing three color frequencies only) resulting in an efficient substitute of the tremendous amount of information needed to describe *perception wise equivalent colors* being represented and created by hundreds of different frequencies.

The research on understanding and on an efficiently compressed representation of (perception wise equivalent) haptic/tactile signals is still in its infancy. Nonetheless in order to pursue a reasonably successful data compression with respect to haptic/tactile signals we still need to come up with realistic and reasonably simple models telling how complex haptic/tactile signals may be efficiently replaced by simpler ones whose description and generation is shorter and more convenient. A substantial amount of research will be needed during the next decade in order to reach the aforementioned simplifications being important for an appropriate data compression of haptic/tactile data. (Some initial research related to the problem at hand is contained in [29], dealing with the perception of roughness).

Comprehensive Model Space for Haptic/Tactile Signals is Missing

The preceding paragraph dealing with haptic/tactile signals has not mentioned yet that there is still a basic foundation missing in the haptic/tactile VR– arena. While we have comprehensive well defined *model spaces* describing audio/visual signals we have nothing really equivalent with respect to haptic/tactile signals. In order to expose two humans at different locations to equivalent sensations via transmitting haptic/tactile signals we need proprietary specialized hardware at both locations, possibly customized for the respective hand/skin and body features of the respective persons. *A basic major step still missing in VR and Multimedia Research including systematically the haptic/tactile sense would be the development of a comprehensive psychophysical model. That would be useful* for describing and generating a well defined vector space or a manifold of haptic/tactile signals generating a comprehensive scalable collection of haptic/tactile sensations. (For the respective audio/visual signals those model spaces are obviously available e.g. for grey value images the vector space of grey value height functions). Furthermore, the development of a-modal and cross-modal layered representations for haptic/tactile, auditory, and visual perception is one of the important challenges in this field.

State of the Art of Haptic VR –Simulation systems

What the state of the art research has done so far is developing VR systems that generate mechanical impacts on the human skin on specified contact areas. Extending this research in a systematic way, we may firstly focus on haptic/tactile VR action and perception of some restricted anatomical parts of the human body e.g. starting with the hands. Using a generic hand model (that may be adjustable i.e. transformable to individual hands based on scanned anatomic data) one could define a gross mechanical impact (force) distribution i.e. a time dependent vector field on specified time dependent contact areas of the hand's skin. (A more sophisticated model should include tactile signals, or even thermal energy distribution) For the description of the mechanical interaction of the hand with other objects, we need a generic time dependent model surface H(t) for the hand together with a time dependent subsurface HC(t) describing the updated contact area with a time dependent Vector force field Vhc(t) assigning a vector to every point in HC(t).

Let us assume one wants to model a real-time haptic/tactile interaction of a human's hand with a deformable object or a textile. Then it would be necessary to send the physical model data of the deformable object to a computer where the person is located. We also need software appropriate for modelling the real-time mechanical interaction of the person's hand with the object controlled by the updated voluntary reaction force of the person's hand. This software system describing an (interaction dependent) haptical renderer would generate signals communicating with the control system of a haptic (hardware) device creating the mechanical force feedback substituting for the interaction of the real object with the human hand.

Organizing the Data Update in a distributed visual/haptic VR-system with Local Clients and a Global Scene Sever

We come now back to the more general initially described scenario where persons at different locations are haptically (and also simultaneously) interacting with each other with a virtual piece of textile. For this scenario one could make the following arrangements. All participating persons have hardware (i.e. force feedback systems) and simulation systems (local clients) running software being capable to simulate scenarios with mechanical interaction of a person's hand with flexible objects like textiles and also with the hands of other humans at their local physical places. Here it would be desirable that this simulation could rely as much as possible on static information that could be provided prior to the mechanical interaction. This (static) information would not change during the mechanical interaction. Therefore it would include fixed physical material parameters of all objects needed for an efficient (possibly specialized) model useful for the simulation of the mechanical interaction. This physical simulation should provide an update of positions of material points and an update of the respective forces being involved (especially in contact areas of mechanical objects.) The latter update would need an update of dynamic parameters resulting from the updated mechanical interactions of all active participants of the virtual scene. It would be desirable to simplify the local simulation model such that the number of dynamic parameters controlling the mechanical interaction would have to be small, so as to keep the amount of (update) data to be transmitted during the mechanical interaction very limited. The update of all the dynamic data of all local clients should be done by a global scene server, responsible for periodical "synchronization" of all local clients in a generalized way. "Generalized synchronization" means here that all local clients *describe simultaneously the same physical state and identical interaction scenes.* To achieve generalized synchronization, a highly precise clock wise synchronization of all local clients and the server is needed. This is possible via modern techniques.

The virtual haptic interaction scenario described above may in principal include large augmented VR –Scenes with virtual as well as real objects being in interaction. This case is in principal a sub-case of the one discussed above where we may additionally assume that one of the human participants together with additional physical objects is partial to the scene. Then what we said above remains still valid: In case real objects including human beings will participate in the interaction one must make sure that the real physical objects provide updates of their dynamic physical parameters to the global scene server.

Distributed visual/haptic VR –Systems without local clients will not work

Concerning the transmission of large audio/visual/haptic databases through the Internet for the purpose of sharing multimodal Virtual experiences the transmission of the relevant data should be organized as described above in order to assure the most realistic experience. Any other design of the system e.g. where all simulations would be done on a central server sending to and obtaining signals from the local hardware systems of the participants will be disadvantageous. This is due to the significant time delay stemming from the communication of the hardware control with a possibly geometrically distant central server. Even the fastest data high way could not eliminate a distance of 10.000 km resulting in noticeable delay and latency time of more than 1/30 second. Especially for the rapid haptic loop needing refreshing rates of 1/1000 second the latter delay is unacceptable.

This implies that the transmission of large audio/visual/haptic databases has to be split in an initialization part, supplying all involving participants with appropriate software and static data needed in the simulation model, including the shared multimodal representations, the abstractions and multi-layered representation of a-modal and cross-modal data and processes. The dynamic update should not rely on many parameters. Clearly, a fast data high way will allow the handling of larger scenes with more participants. It still remains that designing sophisticated models controlling the dynamic change of apparently complex scenes will still be essential. (The latter problem has been dealt with in a comparatively simple special case by the talking head community dealing with audio/visual telephoning. In that area, there have been animation models of faces developed that allow to control relevant dynamic changes of facial expressions practically by a few dynamic parameters only.) After all the key problem for multimodal data transmission of large visual/haptic databases will remain to develop appropriate and mature models that can be used well for the dynamic update of haptic/tactile interaction scenes. In case one would be interested in transmitting data for passive reception of haptic/tactile contact only then it should be possible to describe the respective haptic signal in a normal way using a time variable force field defined on a normalized model hand. Those vector fields may be considered as being a vector space. It would not be too difficult to approximate the latter data in efficient ways.

As far as we are concerned with databases for audio/visual data the classical approaches during the recent 20 years were quite fertile. Clearly, at least with respect to the transmission of visual scenes and advanced stereo audio data it is obvious that partitioning scenes and volume data geometrically in different windows or volumes so as to use parallel transmission of different parts may be helpful at times. Certainly, the transmission of volume data may cause the hardest transmission problems among the ones being mentioned. Here, we have currently problems where it is still impossible to transmit high resolution volume data via the Internet. These transmission problems for volume data will remain being difficult and important. Clearly, ongoing and future research with respect to multi-resolution and wavelet methods will provide promising solutions for those problems. Beyond that, we may need new dedicated hierarchical compression models employing more specific information on how the respective 3D-Volume–structures have evolved and developed.

5.2.1.5 Enhancing virtual human, robotics behaviour, realism and motion

Let us take an example again. We are in the old town of Geneva, and wish to experience the life during the middle-age era. We put our see-through glasses on and witness the ancient merchants speaking Latin and proposing us several merchandises. As we have passed the era of laptops, the computing facilities and memory and network media will be embedded into the walls of each building. This means that walking on, we can see the merchants, discuss/chat with them while seeing the buildings of today of the old town. Thanks to disappearing and distributed computers, and high computing facilities everywhere at any time, we are able fully to live a realistic scene in the middle age Geneva in mixed reality mode. Of course, this scene is totally impossible today due to lack of computing power, bandwidth and memory in a single device.

Furthermore, fast access to large 3D/haptic/audiovisual databases is very important in this context, in combination with content based retrieval. Having (imperfect) virtual humans, robotics behaviour, motion etc, one can use an imperfect entity to query a database of more realistic entities to find similar ones and then to enhance the imperfect one using the results of the query in appropriate (statistical) models.

The recent advances in mobile "augmented ubiquitous interfaces" are envisaged to be built for strength to emulate and extend basic biological principles and communication patterns that biologically have stopped evolving in humans. 'Strength' in multimodal interfaces is defined so that people's multisensory perception can achieve new remarkable levels through fusion of different information sources using embedded computing capacity. Such radical research in new fusion patterns between AR, robotics, virtual human simulation, wearables and ubiquitous computing can lead to new ways of unsupervised audiovisual common-cause techniques for perception of coverage, context awareness and improved and augmented visual performance in a variety of tasks and immensely take advantage of such new 'cyborgian' interfaces. The cyborgian (Clynes [27]) mode of interaction manifestoes itself when a human and other external process interact in a manner that does not require conscious thought or effort. A person riding a bicycle is one cyborgian example, since after time the rider operates the machine (bicycle) without conscious thought or effort, and, in some sense, the machine becomes an extension of the wearer's own body. Of course still important challenges lie in all of the above Networked Media related areas (virtual human, robotics behaviour, realism and motion) as highlighted by the recent AR/VR applications [28][30][31] as shown in **Figure 12** below:

- Limited computational resources: based on the size of the device. However currently most mobile devices are transforming from simple communication or dedicated multimedia devices to powerful computational platforms
- Size, weight of wearable, mobile AR/VR systems should not be a burden but as unobtrusive as possible

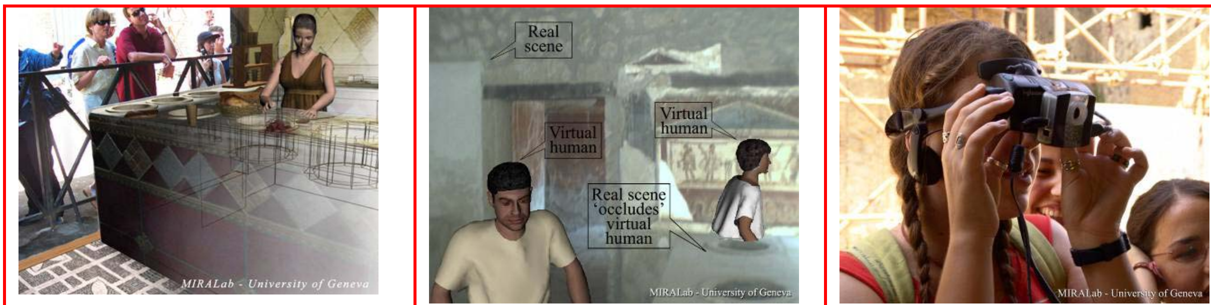


Figure 12: Examples of VR/AR networked media

- Mobile battery life: an important factor of the sustainability of the above AR/VR applications. Except for the smart phone category, most other devices suffer significantly from this aspect, limiting the AR application to a few hours only. Components that have large electrical energy requirements need more batteries, which adds to both weight and size. The power consumption of devices is directly proportional to clock speed and heat dissipation.

- Ruggedness: most mobile AR/VR systems featuring advanced virtual human, robotics simulations with enhanced realism and motion, are early prototypes and depend on the display setup (handheld or HMD) device materials, cables, connectors and cases normally used indoors. The incorporation into everyday clothing as wearables can be proved beneficial as otherwise sensitive electronic equipment needs to be protected from the environment or it will be damaged easily.
- Tracking and registration: these are the basic components of a mobile AR/VR system as specified before. The current trend shows that a combination of tracking modalities are employed for the best results, such as vision or sensor based. This helps to avoid the shortcomings dedicated to each tracking approach, such as lighting conditions or interference.
- 3D graphics and real-time performance: one of the limiting factors for real-time and interactive 3D networked media content has been the absence of dedicated 3D processing units in mobile devices. Nowadays, more of such functionality is incorporated into networked devices as well as the emerging new powerful small-factor PCs (such as Ultra Mobile PC to allow for new possibilities in more compelling media content and applications.
- Networked Media: with the increasing expansion in bandwidth, new breed of audiovisual networked media applications are envisaged and mobile AR/VR systems can profit significantly. However, issues such as content adaptation and sharing, user modelling and personalized interfaces will also need to be addressed for compelling AR/VR applications with advanced virtual human, robotics simulations with enhanced realism and motion.

5.2.2 User Domain and Inclusion

5.2.2.1 Understanding the limits of the applications

Limitations for networked media can also come from the authoring part of such content, from the delivery part as well as from the rendering part. It involves a wide range of technological issues and development tools. Normally, we should work with examples with no limit and identify what are the problems to be solved. This process will in return help to identify the problems better.

Nowadays, the most existing commercial applications using audio/visual multimedia platforms try to take users out of their homes by providing them with small and mobile devices. For mobile games, due to the limitations of device capabilities in terms of small screen, rendering hardware, storage, and battery life, released game titles do not provide as good quality and user experience as those supported by normal desktop or console games even though some of them are developed for 3D. It is a natural trend for users to want more realistic graphics and sounds, so that they feel fully immersed into the provided entertainment application. This is the reason why the next step beyond current game environments will be augmented reality games. Academic researchers already have been trying to apply past popular console-based games to the mixed/augmented reality environments where users play as a game character themselves instead of controlling their avatars. For example, wearing a head-mounted display connected to a mobile computing device, a user plays a DOOM™ style shooting game in which various monsters superimposed over real buildings appear in front of him/her instead of the computer screen. Another example is an augmented PACMAN™ where a user is chased by augmented 3D monsters while he/she searches for 3D items which are in distributed real boxes. Going further than this, the augmented single user game platform could be extended for multi-user environments with the help of broadband wireless networks. With augmented reality devices as an interface, the way of using a portable multimedia player can also be changed such that a user selects music or movie titles shown in a head-mounted display rather than the traditional device screen, which is more comfortable user interface than before. However, from a user's point of view, it requires much more technical evolutions to realize more realistic 3D rendering and easily wearable small display and computing devices with sufficient resolution, computing power, memory, and battery life.

Even if the collection of strategic applications exposed in Section 3.2.1 present new and exciting research challenges, an ordinary approach to their development may not lead *per se* to a revolutionary wave of new services and communication technology. We must be prepared to introduce new elements of innovation in the core applications, introducing such innovative elements into services that at present appear consolidated only for the persisting lack of technological resources. This is even truer when we consider that the connection speeds of data communication *de facto* remove the burdens imposed by the existing communication technology such as handling sophisticated communication protocols and performing data compression. As bandwidth increases, the conventional services might easily become obsolete, and great potential for new developments will be readily available for exploitation.

In the context of Networked Media process, it is then clear that a deep understanding of the limits of each application represents a key factor for the introduction of effective innovation and for the new technology and services to have a real impact on the society. This is an important aspect to tackle, as well-established practice have a natural resistance to innovation, while it is of the utmost importance that to the removal of technological barriers correspond to an equally radical re-thinking of what is offered to the users in terms of services and how these services are organised (let us not forget that many areas of Virtual and Mixed Reality are still largely unexplored and even less has been evaluated). A far-sighted view on application strengths and shortcomings is therefore essential to understand what sort of new enablers should be employed to go beyond what a single component, and in particular the data link component permits.

Proposing new solutions that bring together new elements is then another key aspect in the exploration process of the application limits, and it is expected that the increase of the interdisciplinary aspects in the framework of advanced communication will increase the potential impact of research activities. In fact, cross fertilisation between concepts, algorithms and protocols coming from different research fields, such as 3D graphics, 3D animation and simulation, computer vision, virtual reality, mixed reality, haptics, robotics, digital humans, cognitive and emotional systems, all blended in the range of technologies readily available to develop highly hybrid applications has then the potential to bring into the picture additional innovation and technological advancements.

A final challenge in working with a new generation of applications is that scientists and researchers are confronted with a rapidly moving target in terms of the capabilities of the information technology, and there will not be enough time for the domain to be fully understood and stable. For this reason, charting the strengths and limits of the communication technology, long before it is ubiquitous in the form of applications, may be vital for the Networked Media research as a field. In fact, it is expected that any technology inserted in the picture will be used by many researchers with many different motivations, leading to the result that each technological component will be used in ways not originally intended. Clearly, this has to be added to the fact that any application should cover a necessary portion of the domain space for it to be useful.

5.2.2.2 Cultural diversity/values/differences in media

Services and content provided with networked media should follow user preferences in order to cover all cultural diversities. We could address these differences having social and psychological researchers linked to our research as cultural difference can not be coped only with technical aspects. Then, it could be possible to develop an adaptable application which changes its configuration to fit different cultural and behavioural patterns of a user. On the contrary to most current situations, a user using a strange application has to get used to it no matter what his/her mother tongue, skill, and individual preferences are.

A typical example which shows such a cultural diversity is the language. Researchers have been trying to provide foreigners with an international language translation support in media services such as language tools in web sites and subtitles in movies, etc. However, current effort still has limitations of correct translation and the number of supported languages. Simple syntactic translation can not deliver what an original language exactly means, as sometimes it could include semantic, cultural, and contextual implication. For movie subtitles shown in bottom part of the screen in cinemas, the major problem is that the screen limits the number of available subtitles simultaneously.

To overcome such a problem of multi-lingual aspects, if for example, we use “second life”, we can propose virtual humans as personal digital assistants with personality and cultural models for each person. The virtual human will assist a user so that he/she explains hidden implication of the translated sentences or show suitable subtitles displayed on individual see-through glasses, in order that different audience watching same screen can see a movie with their own preferences. In addition, such digital virtual assistants could also support for individual diversity by collecting and analyzing user preferences for providing better customized media services.

5.2.2.3 User-domain and application inclusion

Proliferation of platforms, technologies, content, applications and services in the new media world, has not been followed by equivalent user immersion and user media experience. The user consumption is based on rich media consumption where the interactivity is an added-value but not a fundamental part of the user experience with such media assets.

User media experience is evolving from the concept of "passive interactivity" to real "user centric media experience" where the active participation of the user is key to the business models of different players in the media chain (e.g. improving the interaction of the user with the content created by the content provider) and in the context of consuming content and services.

Paradigms in the new media world are day by day following a user centric approach where the user domain is getting more and more important due to the change of user roles in media creation and distribution on one hand, and due to the need of more immersive, creative, communicative, collaborative and smart media technologies to address users' desires on the other.

The new applications need to fill the gap between simple content consumption and media supporting true feeling of new experiences. Some of the most important pillars that will need to be addressed in the forthcoming years in the user-domain are:

- The creation of a new scalable middleware that can support, irrespective of the terminal to consume or create the media assets, the most relevant innovations in the media field such as distribution networks, new user input mechanisms (3D), scalable media, etc. to produce new types of interaction with the user.
- New 3D multimodal technologies for user-to-user and user-to-machine interaction to create user multi-sensorial experiences that can be integrated in the user terminals through mesh wireless networks creating a user personal communicative environment.
- Technologies to support user immersive experiences with high quality media (Ultra HD, 3D-HD and beyond) with an end-to-end approach including interactive mechanisms.
- Technologies to measure and track the user experiences with new media such as audience measurement research extended to new convergent media applications and services, and methods to feed this measurement back into design stage to optimise the experience.

- Technologies to create cooperative and participative media, with convergent P2P and M2M (many-to-many) heterogeneous networks for a ranging of terminals.

It cannot be forgotten that technologies such as 3D ecosystems, haptics, 3D user avatars, robotics, etc. are useful for the inclusion of people with disabilities allowing seamless communication with people without disabilities and helping inclusion in the new media world.

Broadening the scope of the applications or the technologies foreseen to answer the issues that stem from networked media could increase the interest of other research communities. This is very easy as every situation in life is interdisciplinary by essence. We should take some case studies in real life as a starting point for each new research proposal, and then include the other disciplines necessary in the case study.

The above mentioned research areas tend to be focused on one usage scenario - namely media sharing. There are however other usage scenarios possible:

Medicine - in the medicine usage scenarios, it is foreseen to allow sharing and effectively searching of the recordings from medical investigations. Such possibility would be an aid for education and diagnosis. The source of the recording could be an endoscopic examination such as bronchoscopy or gastroscopy.

Art - in the artistic scope, there is a need for a dissemination system or art which would allow distributing and effectively searching in the centralized or distributed databases of multimedia artwork.

5.2.2.4 Trying to get people from various disciplines involved

Interdisciplinary – role or sound and emotion in Networked Media

Sound has an important role in our life: "embodied sound media" can contribute to the improvement of novel A/V services (e.g. experience-centric and participative music applications, such as future active listening applications) and support for applications for social (e.g. health) and industrial use cases. The role of sound and interaction span from the auditory channel in multimodal interfaces, to sonic design, to future networked media experience-centric participative applications (e.g. [25] etc.

In 2020, sound and music will be everywhere and for all. The interaction will be much more sophisticated and it will help the engagement of people with music and among themselves. Along with sensors, actuators, microprocessor and wireless gigabit connection facilities embedded into everyday objects, sound will also be a major information display.

Another important aspect emerging from such growing networked media scenario is that we are all annoyed by obtrusive and unwanted bells, whistles and senseless ringtones. Future networked media also concerns the awareness of acoustic ecology and related technology developments, so that not only will the technologies to support A/V content production and distribution be addressed, but also tackling the A/V pollution will be a significant research topic. By bridging the gap between (audio, music, and visual) content and sense, we will be able to take an advantage of sound and music in a more intelligent and engaging way while reducing acoustic pollution to a minimum. The role of sound and music computing is also addressed in the recently launched EU Roadmap on "Sound and Music Computing" [26].

A shift of paradigm on the role of emotion research in future Networked Media

There are initiatives (e.g. W3C, MPEG) to code emotional cues, to define use cases significant for the future scenarios of networked media. They are based on the current paradigm of research on emotions, emotional labels and explicit modeling of affective states.

It is suggested here a widening of paradigm, aimed at empowering the role of emotional interaction in future networked media scenarios and social networks. This consists of facing the modeling of non-verbal subtle communication channels, i.e. strongly related to empathy, resonance, synchronization, kansei [23] For example, it concerns the emotions and empathy communication between two or more humans – not knowing each other – standing still in a waiting room (waiting for a medical treatment, a generic service etc.), and even without exchanging a single word, they exchange a lot of information (emotional, empathic, etc.) while waiting. This subtle as well as powerful aspect of empathy and emotion research should be tackled, modeled, and exploited in future research on networked media (contributing also to social networks issues, ethical issues etc) toward more "sensible" (and therefore effective in experience-centric and participation of users) networked media applications, as well as for networked media infrastructures for health, elderly, and in general social utility.

5.2.2.5 Framework where users seamlessly address technology

An ideal framework would allow for transparently deliver services and contents to users whatever client device they are using and wherever they are located. It could be, for instance, based on dynamic context-driven scalable mechanisms. There is still too much know-how missing and there still exist many constraints for such a framework. One example could be the mobile phone. It is relatively easy to use and most people in a click of a button can operate it. Despite the fact that the mobile is in some way a high-tech device, we do not have the feeling we deal with such a sophisticated technology.

Internet deployment and the services for phone/Internet/TV

In practice, the Internet deployment and the services for phone/Internet/TV are parts of the framework that is driving the high bandwidth of networks today. Following this line of thought, it seems that the realistic targets for the future networked media are the clients (the network will continue to be upgraded as need be and the upgraded portions will also have a lot of dark fiber, waiting for new services). In that respect, there should be effort and technology developed to enable clients (consumers) to “develop” and share their content under a common framework. Thus, clients should be able to create content, process and store it in real-time and with necessary security. For this, one significant direction is the development of framework tools with user-friendly interfaces for the clients (this includes, for example, tools for creating, processing/editing and storing VR content). Among these technologies, which have significant effect on user experience (because they address content directly), it should be made clear that security and DRM are important technologies for such a framework to address a seamless technology. As the amount of content increases, clients will want to be able to protect information that is sensitive and/or personal. The protection is necessary for both storage and transmission purposes, and it comes at different levels (e.g., a user may want to share content as long as he/she gets the credit for it. This can be achieved through watermarking, a DRM technology, and thus there is need to provide effective watermarking on user content easily).

Security

Security is an important technology that needs to be included in the design of all systems, including hardware and software, from the beginning in order to avoid system weaknesses (a common case when security is simply added on at a later stage of development).In addition to security, it is necessary to stretch the need for development of systems that provide specialized (hardware) support for processing multimedia content easily and fast. This is also an important issue, as hardware is a significant component that differentiates performance and cost of services, and is not adequately addressed in Europe. This is the reason that there are no competitive products, especially for consumers, in the EU. Specialized network adapters (for servers and clients) and switches/multiplexers that can process multimedia streams are necessary for delivering services to a large consumer population effectively.

5.3 The challenge of distributed control

"The challenge of distributed control" is a new situation in which neither the infrastructure, nor the service, is controlled by a single entity to deliver a specific service. The following **Figure 13** illustrates a scenario that represents the challenge in networking infrastructures and services under CONTENT NoE question. Rather, the infrastructures and services are based on the composition and interaction of many subsystems and entities, relaying in self-organization and self-interaction technologies. This question has been analyzed in the Networked Media Task Force under two different points of view: the point of view of the infrastructures, and the point of view of the content. Both points of view are detailed in the following two subsections.

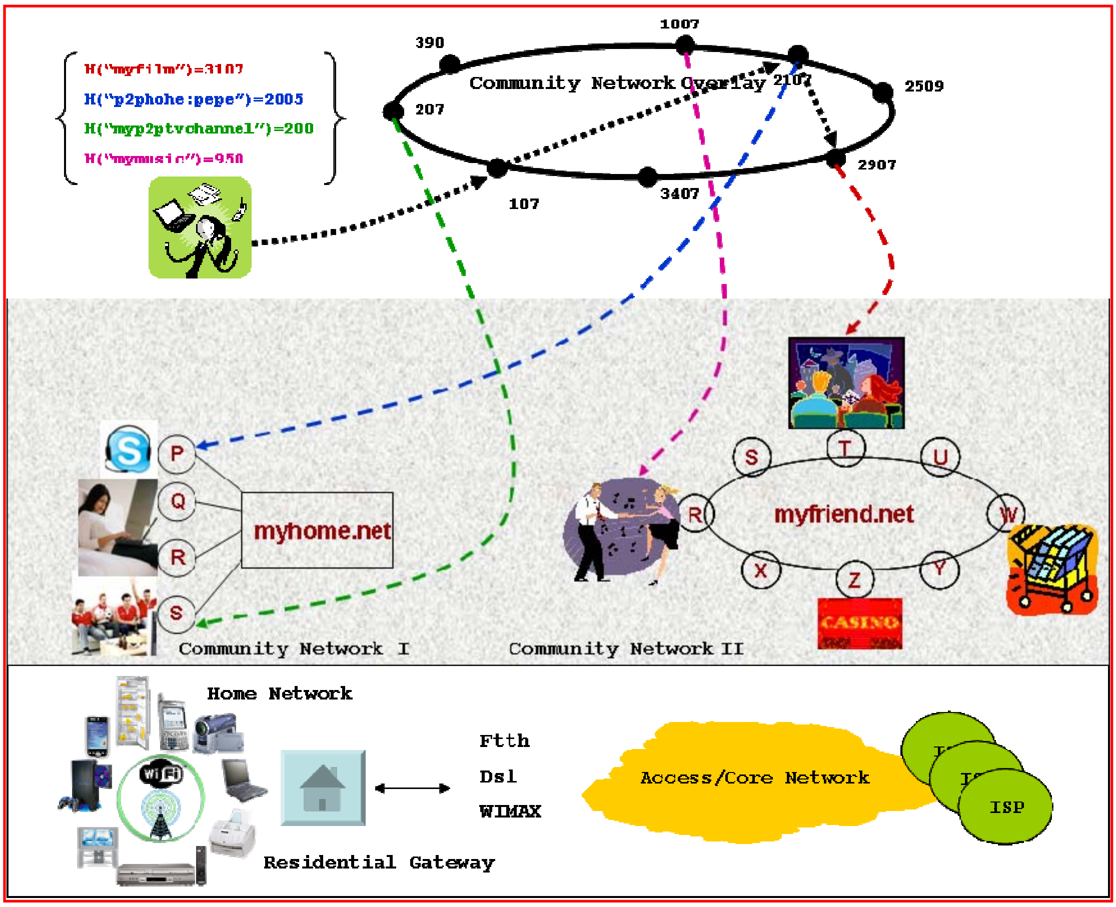


Figure 13: CONTENT NoE scenario: The challenge of distributed control

5.3.1 Networked Media Infrastructures

Future networked media infrastructures will be based on a plethora of different technologies while at the same time have to support the entire range of content types, formats and delivery modes. In such an environment the user's quality demands have to be satisfied optimally. The variety of audio-visual (AV) capable devices is increasing and ranges from High-Definition TV sets and PCs with large high resolution monitors to PDAs and even mobile phones. Most of these devices can be connected to some kind of communication network, like cable, GSM, WiFi, Bluetooth, or power line transmission. Taking this into account a number of challenges have been identified at the first NM-TF workshop. These are namely:

- Dynamic adaptation of infrastructures
- Managing complexity of infrastructures
- Overlay challenges (search, caching, organisation, management)
- Infrastructure management & protection
- Multiple domain multicasting
- Distribution modes, stored vs. distributed
- P2P system as distributed storage system
- System of incentives to become "super node" against different levels of involvement
- Bandwidth symmetry

These are related to the properties of future networked media infrastructures and the way they can adapt to the dynamic system and user behaviour. This indicates that such relatively complex infrastructures will in future have to be organised according to autonomic and self-organising principles in order to adapt to the new challenges of future networked infrastructures.

More specifically the challenges lay in three areas: in the actual network and the integration of different networking technologies, in the overlay networks that will provide end-to-end content delivery, and in the content services infrastructures to optimise the content distribution and end-user experience. Between the three different areas are many important interdependencies and cross-layer issues: besides the well-known problem of mapping non-functional requirements from the application to content services, overlay networks, and networking infrastructure it is necessary to: (a) avoid redundant functionality (b) enable overlays and content services to be context-aware, resource-aware, and location-aware; and (c) to coordinate adaptations that are performed independently in the different planes (such as routing in the community network and routing in the overlay).

The goal is to design an ***Integrated Content Services Infrastructure*** comprising *content delivery* and *content services* within one unifying framework. This infrastructure should provide **seamless, context-aware, end-to-end content delivery**

in a heterogeneous environment in a community context. In the following, the research challenges related to the three different areas are identified.

5.3.2 Networking Issues in Networked Media

A number of research challenges have been identified in the networking domain. Especially the support of mobility in heterogeneous environments is an open issue. More specifically multi-homed mobile nodes and Mobile IP are relevant research topics besides issues such as the impact of the upcoming standards such as IEEE 802.21 and IETF SHIM6 on mobility and on its adaptation at different layers.

Regarding MAC protocols it can be observed that protocols such as 802.11 are not suitable for multimedia content distribution since they still do not have a complete solution to offer QoS guarantees. Hence, QoS provision in the MAC layer is still a relevant research topic that also needs to be pursued in future. Technologies such as IEEE 802.11 and IEEE 802.16 are key technologies for content networks, and they still pose several challenges in terms of optimizing the mapping between underlying quality of service mechanisms and IP-level quality of service approaches. Enhancements to traffic control functions, signalling mechanisms, mobility mechanisms and quality of service management and mapping are the key research topics to be addressed.

Another issue that has to be addressed is misbehaviour detection in heterogeneous and autonomously organised networks. Nodes belonging to such networks may become greedy when competing for resources and behave maliciously in order to increase their own benefits. It has to be researched how nodes can organise themselves autonomously while still providing a resilient and secure infrastructure.

New transmission techniques such as Power Line Communications also have to be taken into consideration. This channel is dissimilar to any other one, and it is very difficult to model and measure. This has to be researched on a theoretical as well as practical level in order to explore its full potential for multimedia services.

Finally, wireless mesh networks are attractive for interactive content exchange. However, the use of wireless links causes communications among nodes to suffer from noise and interference problems. This can be alleviated by the use of non-interfering frequency channels or by endowing each node with multiple radios, or other mitigation techniques recommended by ECC/CEPT technical reports concerning wireless or mobile services. However, research is still required in the area of channel assignment and routing issues, link selection in order to improve performance, optimizing the network for content distribution, efficient gateway selection and integration of wireless sensor networks.

5.3.3 Overlay Issues in Networked Media

The main research challenges in ***Overlay Networks*** are the creation of overlay network infrastructures to support the provisioning of AV services to end-user communities with little or no support from centralized computing facilities. The existence and reliability of such an infrastructure will mainly depend on its members, and on the self-configuration and autonomicity characteristics of robust peer-to-peer networks.

Overlay networks in the context of networked media have to suitably support the provision of AV content services to user communities. The creation of an overlay is mainly justified by the necessity of creating an infrastructure to support the provision of these services among the community members without relying on centralised computing facilities. As such, designing the end-to-end infrastructure involves the creation of new supporting functions that can be implemented either in user terminals or in access networking facilities (e.g. the so called 'home platform'). However, since AV traffic is of great importance for Internet Service Providers, and due to the revenue-generating capability of AV Content services research work has to take into account the role of Telecom operators. In particular it has to be investigated how Telecom operators and Internet Service Providers can exploit Traffic Engineering techniques to match their own management needs with the user's expectations in terms of Quality of Service.

Designing a peer-to-peer infrastructure for networked media is a challenging task, due to the peculiarities of multimedia traffic flows and the dynamicity of community membership. The use of trust and incentives in the organization of peers therefore has to be addressed in this context. Research in this area has to take into account the user level as well as system issues.

Some of the issues related to overlay networks have a wider impact and span, in fact, multiple areas. For instance the specification and measurement of QoS parameters and other metrics that can be used to (1) assess the underlying communication technologies, (2) organize efficient and resilient overlays and (3) suggest most suitable service instances for the end-user should be co-ordinated at the overlay layer. A possible approach to deal with these cross-layer issues is to gather information from the underlying networks and combine it with the higher-level quality assessment and requirements of applications to adjust the overlay networks. In this sense, the use of overlay network technologies is the instrument that makes it possible for user communities to exert some form of control over Quality of Service when no direct interaction with the network is allowed. However, considerable research in this area is still required to come up with an optimal strategy

5.3.4 Content Service Infrastructures in Networked Media

Content Services reside along side and are supported through the underlying communication and overlay infrastructures. They complement and build on their more delivery-focused activities. The main research aspects in this context are related to architecture and service framework issues, service interaction and the integration of service instances within the overall framework. The challenge for the architecture and service framework is to devise a generic model that can accommodate different service types and allow all kinds of interaction modes between autonomous and heterogeneous service entities. This is also reflected in the challenges facing service description and discovery in a completely distributed, heterogeneous environment based on interacting autonomous service entities. Service instances in this context provide examples of the kind and types of services that have to be integrated within the envisaged content network architecture. The main research challenges within this area are:

Co-ordination and management of autonomous services within a unifying content network architecture

Description semantic and service discovery for complex autonomous services in a distributed infrastructure

Service representation and integration within a unifying content network architecture

5.3.5 Research Outlook and Roadmap

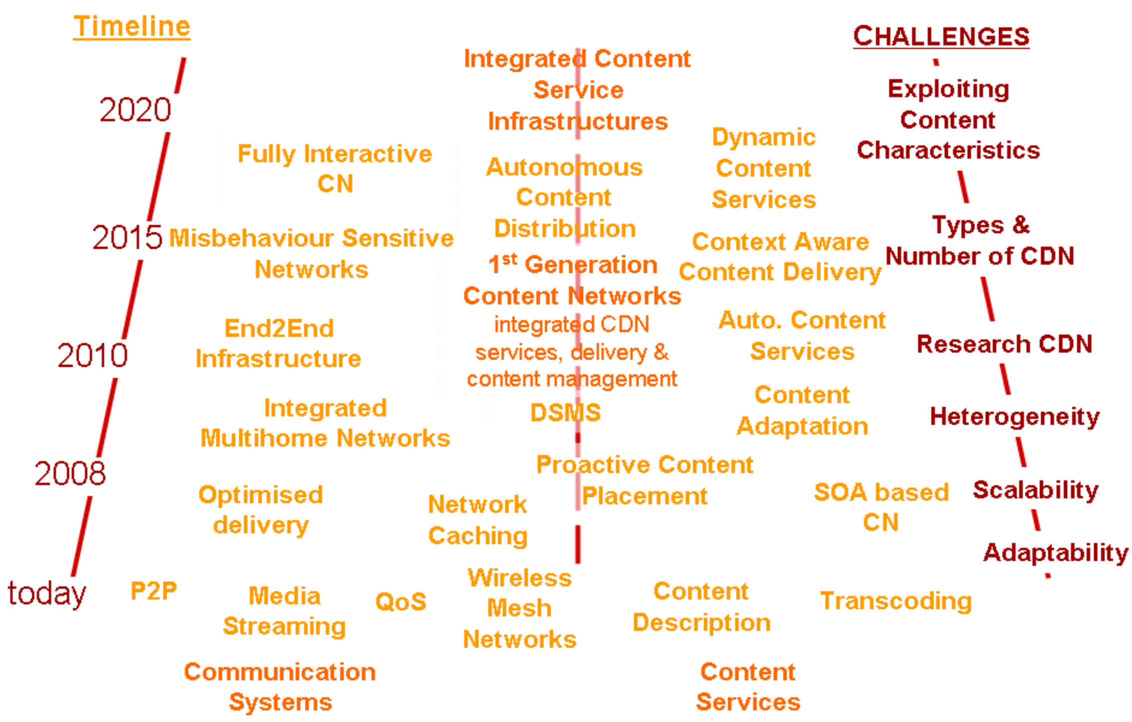


Figure 14: Networked Media Infrastructure Roadmap

Based on the above discussion and earlier results a roadmap has been devised showing the expected development and relevant research areas in the area of

Networked Media Infrastructures in the CONTENT NoE [32]. In the roadmap (as shown in **Figure 14**) we envisage that by 2020 an Integrated Content Services Infrastructure is feasible, comprising content delivery and content services within one unifying framework. The goal is to provide seamless, context-aware, end-to-end content delivery in a heterogeneous networked media environment. Some of the important research problems and issues that have to be addressed on the way to this goal include multi-home networks, misbehaviour-sensitive networks, autonomous content distribution in the communication system and the application of service oriented principles for autonomic content services in Integrated Content Service Infrastructures.

The challenges faced in the context of development towards Integrated Content Services Infrastructure are related to non-functional issues such as adaptability, scalability and heterogeneity. Further, open questions are related to the types and number of CDN and how research CDN can contribute to the development in this domain. The former refers also to the question if ultimately there will be one unifying generic architecture or if there will be multiple specialised CDN architectures as it is the case today (e.g. with television, radio and internet based content delivery networks.) Different solutions for the actual delivery even exist in Internet-based CDN (e.g. P2P streaming, download, VoD, etc.) It is still a major research question if these solutions can be subsumed under one unifying framework or if it would be more efficient to have specific (though loosely coupled) CDN for each. The idea of content networks is based on the assumption that a better service can be provided by taking into account the special characteristics of content. How and to which extend they can be exploited within a content service infrastructure is one of the major research questions.

During the work on this document it has been noticed that it is essential to see the different research topics not in isolation but in an overall context. While the problem domain is structured into layers and planes, cross-layer issues are of central importance in order to tackle the future research challenges in a coherent way and come to real and feasible solutions. This insight has to be considered to achieve good integration of research, because by combining the expertise of the different actors it can be achieved considerably more in the research area of networked media, content networks and services for content networks and home users as individual research activities can.

5.3.6 Interactive, scalable, multi-modal content search

This subsection analyzes the content dimension of “The challenge of Distributed Control”. The content dimension implies all the aspects related with the audiovisual content itself, as for example: production, adaptation, combination, curation, transformation, summarization, or annotation of audiovisual content.

5.3.6.1 Interactive, scalable, multi-modal content search

Multimedia content search using multimodal input mechanisms in scalable media is a key challenge for the European research, trying to overtake the powerful technology and the strong market position of the North American technology for search engines.

Meanwhile ongoing research is focused in creating engines for content search, normally in open domains (Internet) following the same orientation as the text based search engines, other opportunities for search technologies in the new media world could not be forgotten.

Semantic media annotation (e.g. combined with object-based coding) in real or not real time (depending on the application) using isolated or a combination of search methods following a common structure or if possible through the creation of a new standard, is one of the challenges to be pursued.

Multimodal search using as input audio or visual information is not a new challenge. Several years ago the same challenge was presented but never a commercial implementation succeeded in the market. This challenge is still valid and new technologies could be useful to overcome this situation. For example, the creation of content using scene information description is a very useful way but content already created not using any information need automatic description mechanisms. Computer vision and image analysis techniques for scene and pattern recognition could help to adapt a classification mechanism in order to produce scene information but a long way needs to be followed before reaching a minimum threshold of 99% of accuracy.

Other interesting challenge is the combination of multimedia based search with pieces of textual information or pre-indexation that can help with on-line search in real time. Semantic based (sometimes multilingual) textual search is one of the technologies that have advanced in the last years. A combination of techniques in the situations where minimal text information is already placed (e.g. a photograph with a title in a web page) can reduce the amount and accuracy of pure AV techniques.

Scalable search is other challenge that will need a lot of attention due to the potential market that is opening. Mobile-to-mobile or mobile-to-computer (server) search using scalable search techniques are opening a wide area for the creation of tailored search engines for the features of the mobile terminal environment.

Other interesting domain is the creation of search engines for professional usage. Broadcasters and media creators and producers work with a closed environment where media is more or less classified and there are potential areas where the search engines could be placed for professional usage.

Besides many multimedia content providers and internet broadcasting studios which also support content search engines, it is already even prevalent for people to record individual multimedia contents and upload them in multimedia hosting web site. It is also possible for users to organize and broadcast their own programs with preferred music, drama, movies, etc. One problem is that as people can easily access lots of contents created by others, they could be misused without appropriate data right management (DRM), because it is difficult to detect all the violation of copyright. That is the reason why suitable DRM is needed for future abundant individual multimedia contents.

As we have lots of contents which are distributed and stored in every personal broadcasting host machines, a challenging issue is how to search for interesting contents based on multimedia semantics. The semantic-based peer-to-peer search could be done by not only using annotation of the content but also analyzing multimedia itself as it will be more required to find semantically similar contents as current approaches for the distributed content search are based on title, description, and annotation which are described by authors, which frequently incurs wrong search results different from what a requestor really wants to search for. For realizing this, it is also needed the intelligent input query process for the system to figure out which content or which part of it a user wants to see in a more detailed way.

The topics mentioned above will evolve the business model of current broadcasting and hosting companies to meet new requirements as they would have to focus more on the management of audio/visual contents distributed in all over the world. Once contents are not managed in a centralized point, it is another challenging issue to determine payment model for users' using their services.

There are several challenges related to the search efficiency and access of multimedia repository. Some of them are listed below:

Diverse search paradigms– Query by Visual Example (QbVE) was one of the earliest visual search paradigms after the most common simple keyword search. After several years of AVSE (Audio-Visual Search Engines) experiences, it appears that QBVE could not meet all use-cases expectations. Hence, what we call "the page zero problem" arises. That means: what should be the starting page of the search? Some approaches were recently introduced such as "mental search", "composition from visual thesaurus"… These approaches offer different ways to start the search even if no example is available. The preliminary investigations are promising but still unusable in real-time search session. More research work is needed for large scale collections and more ergonomic way of query expression.

Scalability of search methods: many of the research results were validated on limited data collection size. Still much work and investigation need to be achieved to ensure the feasibility on large collection that represents most of the time the real use-cases.

Search in distributed repositories – The distributed repositories role in the content distribution mechanisms is getting more significant with the growth of volume of the stored media. The repositories can be distributed on several levels

Distributed databases – in case of distributed databases the centre of gravity of effective search is on creation of middleware interconnecting the databases in a transparent (from the user point of view) manner. Such solutions are present in the case of keyword search, but are seldom and their creation is challenging in case of QbE search.

Peer-to-Peer overlays – Peer-to-Peer (P2P) networks have evolved in recent years to become one of the most popular means of data retrieval. Recently published statistics show that up to 70% of overall network traffic is utilized by P2P applications. Therefore millions of Europeans are already using P2P networks. P2P access techniques have the potential for both quick data and information about data (metadata) access. They are suitable for legal content distribution in educational, medical and commercial applications. However, this growing sector faces the additional challenge that the increasing data volumes require effective and easy search methods. In addition, a now commonly held opinion among researchers and industry is that simple keyword search is becoming obsolete as an effective data access technique and one of the most promising alternative search methods is multimedia content-based.

One of the most challenging problems that remain **crucial** is **the semantic gap.** The Semantic gap occur when differences between user target and the retrieved information by the systems. This is due to the fact that human perception of the multimedia content is much different from the machine perception (low-level content description). To overcome this gap, there are several promising approaches:

Cross-modal search methods: As the research community realizes the semantic gap problem that may occur from mono-modal content-based search methods, more interest and effort are expected for the development of combined media content description.

Interactive search: including the user in the loop is very promising to bridge this semantic gap. In this regard, relevance feedback methods were investigated in the recent years but efficient methods are still lacking. The efficiency could be measured by the iteration number needed to reach the user target.

Interactive visualization and browsing: this is very important issue that impacts directly the effectiveness of a search session. Still large amount of wok is needed to improve the interaction between the user and the system. In this regard, smart methods of content visualization will greatly help the user to identify and reach his target document.

Distributed search system benchmarking – The creation of different, advanced search systems creates a need for the comparison and ranking of different developed system. For this purpose a set of comparative and innovative benchmarks is required. Such benchmarks should take into account the innovative scope of the content based, semantic search.

5.3.6.2 Content summarization

The goal of content summarization is to create a descriptive summary file of a given content. Such summary file, created automatically, contains the most important features of the given multimedia content.

Large volumes of multimedia content can only be easily accessed by the use of rapid and efficient summarization techniques. Their goals are two-fold: One is to capture the essence of the content in a succinct manner and the other is to provide top-down access into the content for browsing. Towards achieving these goals, signal processing and statistical learning tools are used to generate a suitable representation for the content using which summaries can be created. For most content, a structure is imposed during the production of the content in terms of semantic units such as news stories in news video, or photo-reportage in a picture collection. Therefore, a representation that captures the structure of semantic units that constitute the content would be useful. The user can browse the content using abstractions of each of the semantic units. Hence, past work on summarization of scripted content has mainly focussed on coming up with a Table of Contents (ToC) representation. It has shown that the representation units starting from the raw data up to groups of similar data can be detected using unsupervised analysis. However, the highest level representation unit requires more advanced techniques to bridge the gap between semantics and the low/mid level analysis. Three investigation fields need to be explored more in-depth: - cross-media summarization: Few of the existing summarization methods really make use of all the available modalities and there is thus a strong need for efficient cross-modal unsupervised analysis techniques. - Content-adaptive analysis and representation framework: in order to increase the semantic level of the content representation, there is a need for content-specific processing (for example, the use of domain knowledge in the form of supervised audio-visual object detectors that are correlated with the events of interest). Such content-specific processing have to be integrated in general-purpose multimedia summarization techniques to as late a stage as possible. - Interactive summarization: another way to bridge the semantic gap is to allow the user to interactively bring some knowledge through semi-supervised summarization techniques (active learning, relevance feedback, etc).

One important issue in this regard is the **similarity computation** – Similarity computation is performed when there is a need of calculation of a distance between two media files. The distance is calculated upon the extracted features of the media files (descriptors values) and a given algorithm (similarity computation). A research is required in fields of:

Distance metrics – new distance metrics are required for effective high-level distance measurements in the domain of images and for distance measurements in the domain of video and other multimedia

Effectiveness – research has to be performed in order to achieve higher effectiveness in the similarity computation. Higher effectiveness means both accurate measures as well as lower computation power requirements for the process of similarity computation

The similarity computation will be needed for grouping and clustering algorithms to generate similar entities with regard to a given criteria. Several mono-media approaches have been investigated in the recent years. These latter have been successful in some target applications but fails in many others raising the known problem of "semantic gap". The most promising approaches for summarisation are those which combine several modalities (cross-modal) to generate the content grouping and the representative entities of the summary.

5.3.6.3 Content adaptation, aggregation & digestion

The abundance of user generated content has defined the next generation of a networked life where the roles of creating, editing, publishing, and syndicating the content has been assumed not only by the commercial producers, but also by the users and consumers of such content. This has resulted in multimedia content being ubiquitous. However, due to the associated lack of format, terminal and user homogeneity this also caused problems regarding how this content is displayed.

Content adaptation refers to the process of transforming and adapting content to be compatible with different terminal, network, environment, and user requirements. A mobile terminal with a small screen to display video and at most two channels of audio output will have different requirements than a home entertainment system with HDTV and standard multichannel audio, or than a high-end AV terminal with 3D visual display and hundreds of channels of audio, and a haptic interface. Different network conditions will also require adaptation of the content to achieve an acceptable perceptual quality. Similarly, different user requirements may also make it necessary to change the modality of the content. Displaying audiovisual content to users with sensory impairments require the application of efficient transmoding techniques. The ultimate aim of automatic content adaptation is to allow Universal Multimedia Access (UMA) overcoming these problems and their combinations.

The trend over the past years was to develop content to run on dedicated platforms. This approach tends to multiply the quantity of work since the same content had to be developed several times. A concurrent approach is to develop content at a high quality and to automatically scale it to run on specific platforms according to its capabilities. The content is augmented with adaptation tags taking into consideration bandwidth or device capabilities and can therefore be tuned to specific platforms. Standardisation is an important issue for such content augmentation since a common framework has to be at the core of all components. Research has still to take place to ensure the correct adaptation of different models to a variety of platforms. Focusing on 3d content, an example of a model that has to be adapted is cloth simulation. Different models exist, that are a trade-off between accuracy and computation requirements but a common representation has to be created to allow for automatic adaptation.

With the rise of web 2.0, individual content is produced on a daily basis. The responsibility technology producers have is to make this production as easy as possible. Making content creation in almost any context possible is one of the key issues for the future. Furthermore, aggregation of different kinds of content has to be possible. Again, to make it possible, standardisation efforts are a must.

5.3.6.4 New paradigms of interactions with abstract entities

Most modern mobile terminals are more than terminals alone. Many of them provide software based on commercial or open-source operating systems. They provide high quality loudspeaker dipoles, gyroscopic sensors, simple buzzers that can be used as a vibro-tactile interface, and digital cameras. These modalities of input alone can be utilised in devising new ways to interact with abstract entities. An example is a mobile application that will respond to the content of an SMS by changing the display's colour, or another which when the mobile phone is shaken will synthesise sounds of objects in a box and create the tactile feeling of objects hitting the walls of the box denoting the number of SMS messages or e-mails in the user's inbox.

Interactivity is key in any user-domain environment. Several applications are creating new business opportunities based on the interaction between avatars in a virtual world which has been created by real people. This communication with abstract entities is an example of the relevant domain that is opening in the framework of participatory new media. This had a more or less success in the market and a clear high impact in the news and public awareness.

New challenges and opportunities comes from the extension of the idea of virtual multimedia and multimodal representation of people, adapted not only to leisure in computer games but also to collaborative working, virtual meeting, distance learning, e-health, etc. where technologies such as 3D graphical models, voice synthesis, haptics, smell and taste transmission, are key in producing more *humanised* experiences with abstract entities.

Helping to decrease social exclusion and communication with people with disabilities is an application area of this challenges that should bear in mind.

5.3.6.5 Metadata and indexing

Visual content retrieval systems aim first at describing the visual content of images by means of visual descriptors and then at performing search by visual similarity. Each descriptor usually focuses on a visual aspect of the image and should capture faithfully the image content relevant to the considered aspect. Most used visual aspects to describe the image content are global colour, texture and shape. The advantage of such an approach is that relatively good retrieval results can be obtained with quite simple descriptors (colour histograms, Fourier descriptors, edge orientation histograms, shot boundary detection for videos, etc.) Nevertheless, this global approach has drawbacks in terms of retrieval efficiency since it relies on low-level features, usually having low-level semantic content. A wide range of recent image retrieval technologies, and particularly those based on object class recognition, are based on local features. The visual content of an image is not described by a single descriptor but rather by a set of local descriptors describing different regions of the image. This precise description allows reaching better results than global features since small objects of the image can be described separately and can thus be retrieved or learned more easily. The drawback of these approaches is that they produce a significant larger volume of data, from 10 to 100 times larger than global features. Sometimes, the amount of produced metadata for a single image can be as large as the compressed image itself. For such approaches the design of efficient access methods and indexing structures is essential contrary to global features that can often work with a single sequential scan of the data. Moreover there is a need to develop and evaluate high level semantic features that are well suited for describing more relevant visual content. To date, promising scale-invariant recognition results have been obtained with salient region detector or edge-based primitives such as corners or blobs. There is however a need to generalize such features to produce larger, more interpretable vocabularies or ontologies that can be used For interpretation of raw scene analysis data, search and retrieval from and across different customer premises equipment (differing formats, display terminals, interfaces etc) require advanced communications and far richer methods for automatic content extraction and description than todays' content signatures or indexes. Today's indexes are based on a mixture of automatic metadata generated by the computer and human generated metadata (manual textual annotation is better but also more error prone). Nowadays when we refer to metadata, this involves all kinds of generated data on the raw footage, still image, or other content depositories.

6 EXPECTED IMPACT

6.1 Overall impact

The advances that will come in the next decade in Networked Audiovisual Technologies will have a tremendous impact on different socio-economic dimensions in both Europe and the rest of the world.

The overall impact that we foresee is that the new developments in Networked Audiovisual Technologies will fundamentally modify the way citizens and organizations relate among themselves while also affecting the market rules and structure of the media economical sector, as well as opening huge opportunities for new business sectors.

The blog phenomenon and on-line newspapers have already had a significant impact on the way conventional newspapers realise their businesses. The same will apply to content providers as well as radio and television broadcasters soon. Today, we have ADSL at 20 Mb/s, mobile HSDPA at 1.8 Mb/s, and modern apartments in Japan are fitted with their own 1 Gb/s Internet access. This increase in home-bandwidth at affordable prices makes the Internet a truly multimedia capable network. Over the past year, the number of broadband subscribers in the OECD increased 26% from 157 millions in December 2005 to 197 millions in December 2006. This growth increased broadband penetration rate in the OECD from 13.5 in December 2005 to 16.9 subscriptions per 100 inhabitants one year later. European countries have continued their advance with high broadband penetration rates. In December 2006, eight countries (Denmark, the Netherlands, Iceland, Korea, Switzerland, Finland, Norway and Sweden) led the OECD in broadband penetration, each with at least 26 subscribers per 100 inhabitants. The following graph illustrates the broadband subscribers per 100 inhabitants, by technology in OECD countries in December 2006.

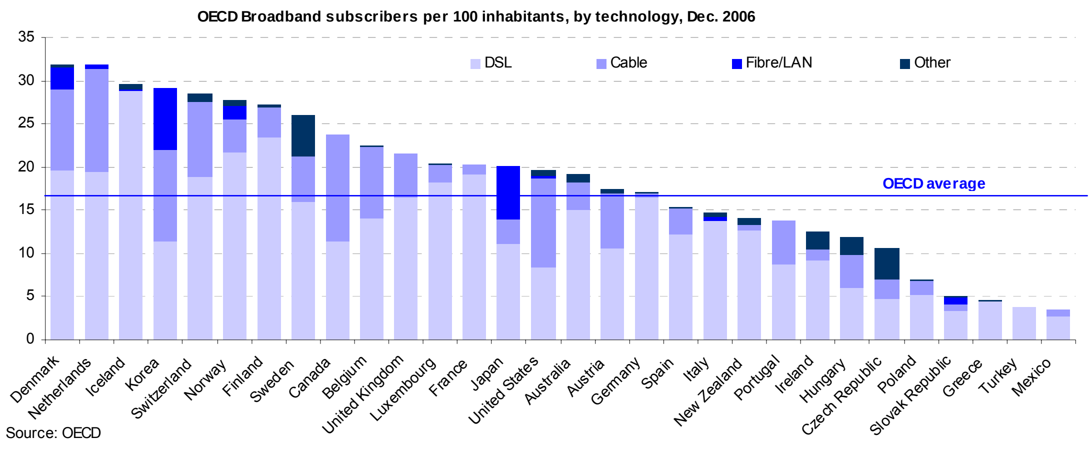


Figure 15: OECD broadband subscribers per 100 inhabitants by technology, Dec 2006, http://www.oecd.org/document/7/LINK_ID:-1

The combination of bandwidth with advanced networking technologies (e.g. peercast, multicast, access-aggregation networks, etc) imply that everyone will have his/her broadcast radio and television station, media production studio, virtual interaction space, and much more, not only at their home, but even carrying with themselves at every moment!

As the Internet is not a receive-only system, like legacy radio and TV, but rather a fully interactive media for everyone, its possibilities to provide new services are just vast. Broadcasters have to fundamentally re-shape their business models, both to adapt to the implications of this new fact and also to be able to take advantage of new business opportunities that these changes will bring. New entrants in the media business, like video-game companies, will certainly pose a menace to the way entertainment will be understood in the coming years. Other new entrants, like telco companies, will also make a huge impact. It is envisioned that a plethora of new integrated services will be widespread:

- Network Games in which groups of users may also talk and see each other while playing.
- Virtual collaboration environments, not just for entertainment (i.e. Second Life), but also as professional virtual workplaces.
- Peer-cast Video broadcasts.
- On-line video clubs.
- Networked wearable computers that provide augmented reality, allowing any-time, anywhere education, entertainment, and work.
- Content distribution services for an integrated audio-visual experience.
- Etc.

Regulators will have to make a significant effort to constantly provide the suitable regulation to promote this trend, while at the same time paying the required attention to privacy, security, equal opportunity, IPR, and other socially relevant aspects.

These dramatic changes have to be anticipated by relevant actors, as they need to begin preparing to have enough time to draw a reaction plan and its implementation. Incumbent companies will have to adapt or die. Administrations will have to provide suitable regulation, or face either a lost opportunity (forbid everything) or business anarchy (no regulation). Users will also have to adapt, at the risk of becoming second-rate citizens, left outside this social revolution.

In the following sections, we analyze the socio-economical impact under different and complementary points of view, leading to an interesting landscape of what the media sector may look like in the coming years.

6.2 Impact on technological areas

The European Union has expanded, and thus represents a vast marketplace now. The potential for companies selling products and services in Europe is enormous. Research and development (R&D) can potentially make a huge impact on the providers and consumers of new audiovisual products and services, especially in the increasingly pervasive content networking environment.

During the past decade, telecommunications and computer networking have essentially merged their R&D activities, and are also increasingly merging the infrastructures and their business sectors. This has led to an extremely competitive and changing sector, where there are opportunities for skilled companies (e.g. Google, Skype, YouTube, etc), and less agile companies may be severely damaged in the process.

On the other hand, networking and media communities have not joined forces yet, and are still essentially separate research (and industrial) communities. Yet, the trend is unstoppable, and the networking and media communities will end-up integrated, whether they like it or not. Incumbents and new-entrants understanding this phenomenon may get huge profits by making advantage of the opportunities that will arise. Incumbent companies, no matter how large and powerful they are now, which do not believe in this trend, do not understand it, or just do not have the ability to re-invent themselves in this new framework, will be driven out of the business, or acquired by a rival or newcomer that has had the ability to adapt.

The potential impact is immense, given the crucial role of networking and the increasing role of content networks and services in supporting so much of society's current – and even more of future – work and leisure.

The technological areas that may be impacted by the future advances include, but are not restricted to:

- Cooperative shared resource assignment and management
- Community networks/services protection

- Mobility and nomadicity in community networks
- Content distribution overlays
- Application scenarios
- Quality metrics
- Content infrastructures
- Service semantics

Dramatic advances in these topics will help to realise the synergy between telecommunications, computer networking, service and content providers, and users. It is clear that in order to achieve the goals of *e-Europe*, audio-visual information has to be provided seamlessly across the entire range of delivery networks to the whole range of possible user platforms.

Due to technological developments, production of multimedia content is no longer restricted to first tier producers. The trend is that end-user communities will be able to efficiently share, distribute, manage, and use audio-visual content.

Future content delivery networks do not only transmit content from a sender to a receiver, but have to be regarded as a content platform that provides support for various content-related tasks and processes. Thus, a content network is an infrastructure that provides a whole range of services to optimize the content experience. Users might be able to access such services for easier navigation, personalised adaptation of content to their needs, etc. In fact, 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, forming a *content network*.

New business areas may arise based on specific aspects associated with services (such as service discovery, service description, service level agreements). These services will be integrated into the overall infrastructure to provide the users with the integrated view based on content service networks.

6.3 Impact analysis by axes

The different environments have tremendous potential for a positive impact on European citizens and institutions. We therefore propose an analysis of their impact taking into account at every moment the relevant aspects from three different axes:

People: *how will these technologies affect the lives of the users, and who will most benefit from their introduction?*

Economy: *how can these technologies drive economical activities, and how well are European economical entities prepared to lead the way?*

Research: *what technological challenges need to be tackled, and what stimulus will the adoption of these environments have on the research efforts?*

The virtual collaboration environment will enable natural communication between people residing in distant locations without incurring the cost and trouble of travelling. The adoption of this technology will depend on the level of fidelity of the communication medium, and the perceived economical/environmental benefits (e.g. lower costs, reduced noise and pollution, etc.) in comparison with face-to-face meetings. This will require people to adapt their behaviour to this new medium, and expect additional tasks as other services (such as document sharing) may be provided in parallel. This facility will primarily attract professionals, but can be modified to suit novice users.

Economically, the use of virtual collaboration environments will result in direct cost savings, and more productivity for dispersed teams. The latter will facilitate pan-European collaboration, and encourage the formation of multi-national teams. On the other hand, the adoption of these technologies will create further opportunities for companies providing supportive services such as content management systems, real-time shared document editing, and visualisation techniques.

Although virtual collaboration technologies have been explored for more than two decades in Europe, their adoption has been slower than anticipated. Therefore, European research should maintain its impetus toward improving such systems by incorporating recent advances of audio/visual communications, developing prototypes, and evaluating how best to accommodate the users.

In recent years, the Internet has re-emerged as a powerful medium for community collaboration, a tool for bringing the contribution of millions and making them matter, and a virtual space where experiences are personalised and shared. By empowering the users, the services created in this manner will serve the community in a better way than if commercial companies intervened in the process. This will allow the free expression of millions of European citizens, the creation of virtual groups across European communities based on common interests, and the promotion of local artistic talents.

Naturally, the power and role of the communication, software and service providers will need to adapt to this paradigm shift. Business models need to be reviewed to incorporate the user-centricity in the value chain. This in turn will impact on how to provide services and charge them, how to respond to users’ queries and manage their rights (if they have contributed to the service creation) and privacy. The reliance on multiple types of providers means that there are many areas that can be improved independently and creatively re-defined, resulting in increased potential returns.

The WWW has originated in Europe, and since its genesis, it has attracted more and more novice users, thanks to the reducing connection costs. This has resulted in a large portion of the population already aware of the possibilities of the Internet, and has started to be actively contributing with very large volumes and varieties of material, such as photos, personal blog, games, music pieces/clips, videos, etc. The wider availability of technological means for communicating and generating content will enable more people to innovate and express themselves. The extents to which this level of empowerment will conflict with the interests of commercial companies remain to be investigated. Other issues relate to the intellectual property and copyrights management of user generated content and services, and privacy concerns for shared content.

The access to educational material via the network and presenting it in media-rich format, will enrich the learning experience, and improve the understanding of the learners. The access to educational material does not have to be limited to the study room, but can be ubiquitous at home and beyond (via portable devices). This implies giving a broader meaning to education that encompasses the traditional meaning and the occasional search for information and details such as manuals, dictionary, medical advice, etc. The availability and accessibility of large volumes of teaching materials will give the European citizens the opportunity to continually improve their knowledge, learn a foreign language, write reviews, etc. Current students/pupils will also benefit from such infrastructures by aligning them with their curriculum and making them multi-lingual.

The constant improvement of the educational system has the direct positive impact on the economic competitiveness of Europe. The development of such systems will rely on a number of economical actors, such as content providers, professional instructors, teachers, software developers, etc. This will stimulate innovation and inter-disciplinary collaboration, which in turn improves the quality of teaching material and its communication.

Europe has introduced different technological-oriented solutions for the improvement of the education system. However, by making more resources available via a user-friendly interface, and accessible from the comfort of the home, researchers will need to investigate innovative ways for collecting, searching and conveying multimedia information, and the provision of varying levels of details depending on the presentation medium and the user's individual requirements.

The ability to access the sheer amount of data stored in different computing devices at home, including pictures, videos and music, will make personal data truly ubiquitous. This will radically change people's consumption patterns of multimedia content. Instead of carrying multiple single-application devices, they will use one multi-purpose device which can stream or acquire the data remotely. Besides, people will be able to access their personal assets and have shared experiences with others.

The commercial stakeholders on the service delivery chain are numerous, hence, each party may provide enhancements on the segment it manages, but they all need to work seamlessly with the each other, for providing the best experience to the user.

Naturally, there are numerous technical challenges to drive the research efforts. The storage and bandwidth capacity at home will always outperform that in a mobile environment. In this context, it is important to find the right balance between quality and resource requirements. Furthermore, security issues are particularly critical in these scenarios because the bulk of the data is transmitted over wireless links.

The provision of cheaper health monitoring systems will extend medical care in Europe to people living in remote areas, the aged population and citizens with physical disabilities, from the comfort of their home. The automated and non-intrusive health monitoring system will appeal to patients, who were previously technology-agnostic, because of the clear benefits.

The implementation of tele-monitoring services requires the integration of numerous economical players, from sensor designers, to medics, to data processors. This in turn will foster the formation of specialised companies and strengthen European competitiveness in health care service provision.

There are still many technical challenges involved in the development of medical tele-monitoring. Scientific research in academic as well as industrial sector will be stimulated to design increasingly reliable and comfortable monitoring devices. Furthermore, research will be needed in the areas of personal data protection, identification of legal roles and establishment of patient-carer contracts in the context of this new form of medical care.

Sports simulation has become an important part of the gaming entertainment industry (judging by the number of sports-based games released in recent years). However, the level of interaction with the game (i.e., input/output modalities) is very limited. By making the game more realistic (in terms of graphics, interaction and behaviour), more people, who are not regular game players, will be interested in participating. Furthermore, gaming can be used by professional sportsmen/women for training and strategy evaluation, especially for dispersed teams. This will provide cheaper, yet effective, means for training for European athletes.

Europe is well positioned in the electronic game industry, and advances in hardware and software will further strengthen its leadership in this area. Furthermore, the availability of these advanced technologies will drive the creation of many companies to develop novel interactive applications.

The increasing realism of games and its potential use as a training medium requires further research in how to present it to professional players and how to evaluate the performance gain from it in the best possible way. Furthermore, the involvement of more people in interactive sports gaming and its potential mix with other media types raises technical challenges and usability issues that are the key areas for further exploration.

6.4 Impact on the value-chain: from research to marketplace

Dissolving technical boundaries

One of the main impacts on the networked media research is dissolving technical boundaries of different disciplines related to the personalized networked media environment. The networked media will have a strong impact on restructuring the research capacities and technological skills in the wireless wearable mobile media field. Today, wireless wearable mobile media is based on a large spectrum of fundamental domains, comprising networking, wireless communication, wearable devices, security, interaction, and multimedia and computer graphics. The evolution of the domains is conditioned by how research fields of these different fundamental domains will be able to communicate. In previous collaborative research experiences, many technical difficulties have been experienced mainly due to the different characteristics in media types with different protocols, a wide range of networks and device capabilities, and different user models which require knowledge on totally different areas such as networking, wearable interface, computer graphics, video, image processing, signal processing, etc. These differences have been preventing the realization of new innovative concepts due to limited expertise in the related areas. To accelerate the emergence of the networked media technology, it is of paramount importance to allow the sharing of these technologies and the reuse of the huge amount of work it represents.

The major impact of networked media research is to integrate these related but partly separated areas into a global vision for new networked media environments. Such a potential technological vision will lead a new way towards the developments in considered areas, which will in turn contribute to restructure the research labs to transform into truly interoperating multidisciplinary entities.

Spreading innovations

The expert knowledge and innovations made in the networked media research will be carefully categorized and presented to public to enhance the efficiency of research activities and to provide broader accessibility on the new concepts and technical innovations. The networked media research will provide international research infrastructures for research activities in the field of wireless wearable mobile media, which will be developed and maintained also for the needs of other communities. This infrastructure and related activities will establish sustaining cooperation among research institutions in Europe.

The networked media research will also have a durable impact on European research in the training of students and young researchers. PhD students and post-doc researchers, highly trained in the multicultural and multidisciplinary context, could trigger research activities at a larger scale. As a result, industry and academia will be brought together for the exchange of experiences and the formation of common understanding on needs and exploitation of results.

Contributions to standards

The networked media will generate a new vision on the nomadic daily life. Development of networked media devices will take new standardization from standalone mechanical device to networking and adaptable devices. On the other hand, to control these adaptable networking devices, a new device interface design, which can be personalized according to the user's requirements, will initiate the future standardization actions on this new field.

Opening a new market

Researching suitable technologies that target networked media oriented environments will generate a new vision for the people to discover the comfort of flexible living while accessing all the related multimedia content and devices around them. In this way, new areas of development will be oriented in marketplace both for hardware devices and relevant software technologies. New business models will be developed and the barriers for development of this technology with its relevant extensions will be reduced.

The networked media research will deliver prerequisites for the commercial exploitation of services combining communication frameworks and multimedia content providing, which are crucial to the achievement of the convergence of various networked media services. Through the commercial exploitation of services combining audio-visual platforms, 3D graphics, and other multimodal interfaces, the networked media research will be able to produce the required market analysis and the business plans in close connection with the real systems and prototypes.

7 CONCLUSIONS

Availability of information at all times and in all places is becoming crucial to the development of modern day society. Information is already available in the form of various media such as audio, video and text. However, as this information becomes increasingly digitized, the potential to transmit them over high bandwidth networks is also ever increasing. It is absolutely essential that Europe be at the forefront of this sweeping change in the manner in which we make information available to society.

In order to ensure this, the Networked Media Task Force (NM-TF) has been established, under the Networked Media Unit. It's purpose is to spearhead the search for new research areas that will ensure that Europe becomes a leader in this very important sphere of scientific and technological advancement. Many ongoing initiatives (such as INTERMEDIA, CONTENT, VISNET II, NEM, NETS, FIND) already represent the major trends in Networked Media research. They are aimed to produce significant advances in many diverse research areas, for e.g., systems and platforms, networked media content creation, content discovery, selection, distribution, consumption.

It is, however, essential that we think ahead. This white paper, written by the NM-TF, is specifically intended for the purpose of identifying future research challenges in Networked Media. We have focussed on three main axes of future progress - true broadband support, personalized media and distributed control.

In order to meet the above challenging goals, the future Networked Media should be designed, developed and released with the following considerations:

- *New form of the Networked Media*, which includes large amount of meta-data and user generated tags so that a system can easily analyze content in a semantic way.
- *Easy way to access the Networked Media*, which will be distributed in different geographic repositories and connected to the users with their heterogeneous (mobile or stationary) devices in various underlying network conditions.
- *Protection of the Networked Media from any illegal and malicious usage*, since all content can be accessed anywhere and at any time, it is essential to provide secure and safe access to private and secret information, lest it should be susceptible to illegal use.
- *Payment models for commercialization of the Networked Media in new business markets*, which will become more complicated than those in the current situation because the Networked Media will be produced, managed and consumed by tightly linked interaction among various kinds of service providers, producers and end-users.

In conclusion, we expect that the Networked Media of the future will have tremendous impact on different socio-economic dimensions in both Europe and the rest of the world in a way that will affect both individual citizens and organizations, will affect market rules and structure, and will open huge opportunities for new business sectors. All of these planned characteristics and advances in technologies must ultimately be of benefit to all users, without any discrimination, in order to increase the quality of life one step further.

8 REFERENCES

[1] INTERMEDIA Deliverable, "D3.1 State-of-the-Art Report on Approaches Toward User-Centric Multimedia", http://intermedia.miralab.unige.ch

[2] INTERMEDIA Deliverable, "D10.1 Media convergence framework for a common platform", http://intermedia.miralab.unige.ch

[3] B. S. Manjunath (Editor), Philippe Salembier (Editor), Thomas Sikora (Editor). Introduction to MPEG-7: Multimedia Content Description Interface. Wiley, 2002.

[4] Thomas Deselaers, Daniel Keysers, Hermann Ney. FIRE -- Flexible Image Retrieval Engine: ImageCLEF 2004 Evaluation. In Multilingual Information Access for Text, Speech and Images. Proceedings of the 5th Workshop of the Cross-Language Evaluation Forum. CLEF 2004, Lecture Notes in Computer Science, pages 688-698, Bath, UK, September 2004.

[5] M. Stemm and R.H. Katz, Measuring and reducing energy consumption of network interfaces in hand-held devices, IEICE Transactions on Fundamentals of Electronics, Communications and Computer Science (August 1997).

[6] M. Zorzi and R. Rao, Energy constrained error control for wireless channels, IEEE Personal Communications (December 1997).

[7] P. Lettieri, C. Fragouli and M.B. Srivastava, Low power error control for wireless links, in: Proc. of ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom), Budapest, Hungary (September 1997) pp. 139–150.

[8] J.M. Rulnick and N. Bambos, Mobile power management for wireless communication networks, Wireless Networks 3(1) (March 1997) 3–14.

[9] A. Chockalingam and M. Zorzi, Energy efficiency of media access protocols for mobile data networks, IEEE Transactions on Communications 46 (November 1998) 1418–1421.

[10] S. Singh and C.S. Raghavendra, PAMAS: Power aware multi-access protocol with signaling for ad hoc networks, ACM Computer Communication Review (July 1998) 5–26.

[11] S. Singh, M. Woo and C.S. Raghavendra, Power-aware routing in mobile ad hoc networks, in: Proc. of ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom), Dallas, TX (October 1998) pp. 181–190.

[12] M. Zorzi (guest ed.), Energy management in personal communications and mobile computing, IEEE Personal Communications 5(3) (June 1998).

[13] C. Perkins, Ed. IP Mobility Support for Ipv4. IETF RFC 3344. August 2002.

[14] A. Valko, "Cellular IP: A New Approach to Internet Host Mobility," ACM SIGMOBILE Computer Comm. Rev. vol. 29, no. 1, pp. 50-65, Jan. 1999.

[15] R. Ramjee, K. Varadhan, L. Salgarelli, S.R. Thuel, S-Y Wand, and T. La-Porta, "HAWAII: A Domain-Based Approach for Supporting Mobility in Wide-Area Wireless Networks," IEEE/ACM Trans. Networking, vol. 10, no. 3, pp. 396-410, June 2002.

[16] A.C. Snoeren and H. Balakrishnan, "An End-to-End Approach to Host Mobility," Proc. MOBICOM, Aug. 2000.

[17] D. Maltz and P. Bhagwat, "MSOCKS: An Architecture for Transport Layer Mobility," Proc. IEEE INFOCOM '98, pp. 1037-1045, Mar.-Apr. 1998.

[18] R. Stewart et al., "Stream Control Transmission Protocol," Internet Engineering Task Force RFC 2960, Oct. 2000.

[19] E. Wedlund and H. Schulzrinne, "Mobility Support Using SIP," Proc. Second ACM/IEEE Int'l Conf. Wireless and Mobile Multimedia (WoWMoM'99), Aug. 1999.

[20] Yi Pan, Meejeong Lee, Jaime Bae Kim, and Tatsuya Suda. An End-to-End Multipath Smooth Handoff Scheme for Stream Media. IEEE Journal on Selected Areas in Communications, Vol. 22, No. 4, May 2004.

[21] Voyiatzis, A. G., Fragopoulos, A. G., & Serpanos, D. N. (2004). Design Issues in Secure Embedded Systems, (chapter) In The Handbook of Embedded Systems, R. Zurawski (ed.), CRC Press.

[22] A. Greenfield, Everyware: The Dawning Age of Ubiquitous Computing, New Riders Press, 2006.

[23] A. Camurri (Editor) (1997) Proceedings Intl Conf on KANSEI – The Technology of Emotion, DIST-Università di Genova, October 1997.

[24] A. Camurri, G. Volpe (Editors) (2004) Gesture-based Communication in Human-Computer Interaction. Selected and revised papers of the 5th International Gesture Workshop (GW2003), Lecture Notes in Artificial Intelligence, LNAI no.2915, Springer-Verlag.

[25] A. Camurri, T. Rikakis (Guest Editors) (2004) Special Issue of IEEE MULTIMEDIA on Multisensory Communication and Experience Through Multimedia, Vol.11, No.3, Jul-Sep 2004, IEEE CS Press.

[26] The EU Roadmap on Sound and Music Computing, http://www.soundandmusiccomputing.org/roadmap

[27] M. Clynes, "Cyborg II: Sentic Space Travel", The Cyborg Handbook, Chris Hables Gray, Steven Mentor, Heidi J. Figueroa-Sarriera (eds), Routledge, October 1995

[28] A. Egges, G. Papagiannakis, N. Magnenat-Thalmann, "Presence and Interaction in Mixed Realities", The Visual Computer, Springer-Verlag Volume 23, Number 5, May, 2007

[29] D. Allerkamp, "A Vibrotactile Approach to Tactile Rendering", The Visual Comp 23(2):97-108,2007

[30] P. Volino, N. Magnenat-Thalmann, "Resolving Surface Collisions through Intersection Contour Minimization", ACM Transactions on Graphics (SIGGRAPH 2006 proceedings), ACM Press, 25(3), pp. 1154-1159. July 2006

[31] H. Kim, C. Joslin, T. Di Giacomo, S. Garchery, N. Magnenat-Thalmann, "Device-based Decision-making for Adaptation of Three-Dimensional Content", The Visual Computer, May 2006, Vol.22, No5, pp.332-345.

[32] CONTENT NoE Deliverable D2.1. CONTENT Research Vision and Roadmap. May 2007. http://www.ist-content.eu

ANNEX I. CONTRIBUTORS LIST

Last name	First name	Affiliation	E-mail
Alvarez	Federico	UPM, ES	fag@gatv.ssr.upm.es
Azcorra	Arturo	University of Carlos III of Madrid	azcorra@it.uc3m.es
Bergamasco	Massimo	Haptics	bergamasco@sssup.it
Boujemaa	Nozha	INRIA	Nozha.Boujemaa@inria.fr
Camarillo	Gonzalo	IETF, Helsinki University of Technology	gonzalo.camarillo@ericsson.com
Camurri	Antonio	Univ. of Genova/DIST, IT	antonio.camurri@unige.it
Chiu	Dah-Ming	China University of Hong-Kong	dmchiu@ie.cuhk.edu.hk
Cisneros	Guillermo	Telecommunications School of the UPM	gcisneros@etsit.upm.es
Dogan	Safak	University of Surrey	S.Dogan@surrey.ac.uk
Donnelly	Willie	Waterford Institute of Technology	w.donnelly@tssg.org
Izquierdo	Ebroul	Queen Mary University London	ebroul.izquierdo@elec.qmul.ac.uk
Firth	Chris	Thales Research and Technology UK	chris.firth@thalesgroup.com
F. Kukielka	Jose	University Carlos III Madrid, ES	Kukielka@it.uc3m.es
Grega	Michał	Univ. of Science & Technol./AGH, PL	grega@kt.agh.edu.pl
Guerrero	Carmen	University of Carlos III of Madrid	guerrero@it.uc3m.es
Ikonomou	Demosthenes	EU	Demosthenes.Ikonomou@cec.eu.int
Izquierdo	Ebroul	Queen Mary Univ. London, UK	ebroul.izquierdo@elec.qmul.ac.uk
Kavassalis	Petros	University of Chios	petros@cfp.mit.edu
Kondoz	Ahmet	University of Surrey	A.Kondoz@surrey.ac.uk
Laso-Ballesteros	Isidro	EU	Isidro.LASO@ec.europa.eu
Leszczuk	Mikolaj	Univ. of Science & Technol./AGH, PL	leszczuk@kt.agh.edu.pl
Lim	Mingyu	Miralab/CUI/University of Geneva	Mingyu.Lim@miralab.unige.ch
Magnenat-Thalmann	Nadia	Miralab/CUI/University of Geneva	thalmann@miralab.unige.ch
Mauthe	Andreas	Lancaster University	andreas@comp.lancs.ac.uk
Misra	Vishal	Columbia University	misra@cs.columbia.edu
Papagiannakis	George	Miralab/CUI/University of Geneva	papagian@miralab.unige.ch
Plagemann	Thomas	Univ. Of Oslo, NO	plageman@ifi.uio.no
Rodríguez-Roselló	Luis	EU	Luis.Rodriguez-Rosello@ec.europa.eu
Romaniak	Piotr	Univ. of Science & Technol./AGH, PL	romaniak@kt.agh.edu.pl
Tekalp	Murat	Koc University Turkey	tekalp@ece.rochester.edu
Triana	Eugenio	Visiting Prof	etrigar@eugeniotriana.com
Weber	Andreas	University of Bonn	weber@cs.uni-bonn.de

Wolter	Franz-Erich	Leibniz Univ. Hannover/Welfenlab, DE	fwolter@mit.edu few@gdv.uni-hannover.de
Wu	Enhua	Macau University	EHWu@umac.mo