



SALERO

Digital Media Technology, Research, and Production Trends Report. Version 4.

SALERO Deliverable D10.3.4



Digital Media Technology, Research, and Production Trends Report. Version 4

SALERO Deliverable D10.3.4

SALERO identifier: SALERO_D10.3.4-DigitalMediaTechnology_Final.doc

Deliverable number: D10.3.4

Author(s) and company: S. Cheung (DLLNI); S. Miarnau, A. Monroig (AM);
R. Fach (DFT); R. Villa (UG), T. Stolt,
M. Tuomola , T. Korpilahti (TAIK); G. Kienast, (JRS);
T. Bürger (LFUI); C. Goodman (PGP),
M. Matthews (BLITZ); A. Evans (FBM-UPF);
O. Mayor (UPF); C. Cullen (DIT); C. Monzo (URL)

Work package / task: WP10

Document status: Final

Confidentiality: Public

Version	Date	Reason of change
1	2009-10-29	Document structure proposed by DLLNI
2	2009-11-26	Version for internal review (pending TAIK input)
3	2009-11-27	Version for internal review
4	2009-12-04	Final version with typo corrections from internal review

The work presented in this document was partially supported by the European Community under the Information Society Technologies (IST) priority of the 6th framework programme for R&D.

This document does not represent the opinion of the European Community, and the European Community is not responsible for any use that might be made of its content.

This document contains material, which is the copyright of certain SALERO consortium parties, and may not be reproduced or copied without permission. All SALERO consortium parties have agreed to full publication of this document. The commercial use of any information contained in this document may require a license from the proprietor of that information.

Neither the SALERO consortium as a whole, nor a certain party of the SALERO consortium warrant that the information contained in this document is capable of use, nor that use of the information is free from risk, and does not accept any liability for loss or damage suffered by any person using this information.

Table of Contents

1	Executive Summary	1
2	Introduction.....	2
2.1	Purpose of this Document	2
2.2	Scope of this Document.....	2
2.3	Status of this Document.....	2
2.4	Related Documents	2
3	Digital Media Technology Today	3
3.1	The Content Sector and Creative Industries.....	3
3.2	Domestic Digital Media Technology.....	3
3.2.1	<i>“The Digital & Connected Living Room”</i>	3
3.2.2	<i>Blu-Ray Adoption</i>	4
3.2.3	<i>Games</i>	5
3.2.4	<i>User-Generated Content</i>	6
3.3	Professional Digital Media Technology.....	6
3.3.1	<i>Introduction</i>	6
3.3.2	<i>Acquisition</i>	7
3.3.3	<i>Storage</i>	8
3.3.4	<i>Post Production</i>	9
3.3.5	<i>Industry Trends</i>	9
3.4	Video & Audio Standards.....	10
3.4.1	<i>Video Standards</i>	10
3.4.2	<i>Speech and Audio Standards</i>	12
3.4.3	<i>Production Standards</i>	13
4	Digital Media Production Trends	15
4.1	Audiovisual Market Production Trends: Cross Media and Movies	15
4.1.1	<i>Movies and 3D</i>	15
4.1.2	<i>Cross-Media and Social Media Production Trends</i>	16
4.1.3	<i>Cross Media Interfaces and Cross Media Research</i>	16
4.2	Games	17
4.3	Television.....	18
4.4	Advertising	23
5	Digital Media Research	24
5.1	Digital Media Research by SALERO Partners.....	24
5.1.1	<i>Character Animation</i>	24
5.1.2	<i>Semantic Modelling, Annotation and Search</i>	24
5.1.3	<i>Audio Transformation</i>	24

5.1.4	<i>Speech in Animation</i>	24
5.1.5	<i>TTS Synthesis</i>	25
5.2	Similar Media Research Projects by External Groups	25
5.2.1	<i>Character Animation/Automatic Programme Generation</i>	25
5.2.2	<i>Semantic Media Annotation</i>	25
5.2.3	<i>Content Search & Retrieval</i>	25
5.3	Emerging Results from SALERO Research	26
5.3.1	<i>Character Animation</i>	26
5.3.2	<i>Media Annotation</i>	26
5.3.3	<i>Voice Transformation</i>	26
5.3.4	<i>Speech Analysis & Annotation</i>	26
5.3.5	<i>Content-Based Retrieval</i>	26
5.3.6	<i>TTS Synthesis</i>	27
5.4	Future Directions for SALERO Research Partners	27
5.4.1	<i>Character Animation/Automatic Programme Generation</i>	27
5.4.2	<i>Media Annotation</i>	27
5.4.3	<i>Voice Transformation</i>	27
5.4.4	<i>Speech Tagging</i>	28
5.4.5	<i>Image/Video Retrieval</i>	28
5.4.6	<i>Text-To-Speech Synthesis</i>	28
6	SWOT Analysis – in the Light of Current Trends	29
7	Glossary	30
8	References	31

1 Executive Summary

This deliverable follows on from the D10.3.3 survey of digital media technology trends, delivered in November 2008, and maintains the same approach to identify trends in the development of intelligent media technologies/services. D10.3.4 was written in November 2009 towards the end of the fourth/final year of the SALERO project, and while many of the trends have developed as predicted, some unpredicted and potentially disruptive technologies have appeared.

Through both the professional and domestic sectors, we continue to see convergence taking place between digital technologies to create the necessary preconditions for Intelligent Content and the funding of a significant body or research designed to add intelligence or smart functionality to media tools, products and services.

In addition to the trends, digital media research carried out within the SALERO project by individual partners and their future research direction in those areas are also outlined. The SWOT analysis on SALERO from D10.3.3 is still largely valid.

2 Introduction

2.1 Purpose of this Document

To exploit the results of SALERO successfully, it is necessary to understand the present and evolving market place. Gathering information on this is time consuming. Currently, while each manufacturer carries out what market awareness studies they can, these are restricted to their individual niche and location within the chain: because of time and cost pressures, each manufacturer tends to look only at his direct competition.

The integration of activities within the consortium provides SALERO with the opportunity to make an economy of scale and carry out this function at a project level to produce a broader picture of intelligent media technology and markets as they evolve, to the benefit of all the partners and as the basis for exploitation planning.

This deliverable is the fourth in a series of reports designed to record and present trends in media production in the sectors that would be related to SALERO outputs (including the film post production area, broadcasting postproduction, games production, 3D animation, online streaming, mobile content production and audio post production) and trends in research that may support the development of intelligent content technologies.

2.2 Scope of this Document

This document concentrates on events and developments occurring in the period from 1st December 2008 to 30th November 2009.

2.3 Status of this Document

This is the final version of D10.3.4.

2.4 Related Documents

Before reading this document it is recommended to be familiar with the following SALERO documents:

- D10.3.1 Digital Media Technology, Research, and Production Trends Report. Version 1.
- D10.3.2 Digital Media Technology, Research, and Production Trends Report. Version 2.
- D10.3.3 Digital Media Technology, Research, and Production Trends Report. Version 3.

3 Digital Media Technology Today

3.1 The Content Sector and Creative Industries

The content sector and creative industries shared mixed fortunes in 2009. As might be expected, the ripple effect of the credit crunch was felt in all areas, with film production in the US leading the way down. In late April of 2009, it was reported that between 75 and 80% of Los Angeles based technicians and film-crew workers were unemployed. For the whole of 2009, a total of 8 'big-budget' movies were scheduled for production, which compares with 21 in 2008 and 71 in 1996. Television and commercial production suffered similarly, primarily as a result of severely restricted investment opportunities. The credit crunch created a situation in the first half of 2009 where raising funds of any magnitude was to all intents and purposes impossible.

The situation in Hollywood was broadly representative of the prospects for the rest of the developed world, and demonstrated how inextricably linked are not only international economies, but also service and creative industries with the funding and banking sector. The universally condemned behaviour of the banking industry in general and financial speculators in particular resulted in the production of new content becoming extremely stifled.

Later in the year, many parts of the global economy moved out of recession as a result of concerted and coordinated stimulus packages. The games industry seemed to be returning to a more upbeat mood towards the end of the year with the launch of titles such as 'Call of Duty: Modern Warfare 2' expected to gross £150 million in its first week. It was predicted by some industry sources that 2009 could be the year in which global games revenue beat that of the film industry worldwide, however whether this would be as a result of games performing well or films performing poorly was less clear. In addition, intense retail competition in yet another underperforming sector further confused the picture. Modern Warfare 2's recommended retail price was set at £55, far in excess of the more usual £39.99 price point for similar games. However even before launch, larger distribution channels were fighting for volume, with Sainsbury's supermarket offering the game for £26. In other content sectors, this push towards volume over margin was also demonstrated, and may be indicative of the increasing gap between large, multi-market channels and smaller more specialised distributors. An example of this would be the launch of Dan Brown's new best-seller 'The Lost Symbol', which was launched with a recommended retail price of £24.46 and yet could be bought from the supermarket chain Tesco for £7.99.

Seemingly, the good news for the content and creative industries is that many major economies (with the notable exception of the UK) are either out or moving out of recession as the year draws to a close. However, this move into growth comes primarily as a result of unprecedented stimulus by governments, which will need to be repaid at some point. Further, this level of stimulus is in itself confusing the issue of economic performance as it is as yet unclear whether the US economy in particular has returned to a sufficiently resilient to continue to grow once the support of such stimuli is removed.

The most notable takeaway from 2009 according to Morgan Stanley was that the mobile internet is and will be bigger than most think. Apple mobile shares are expected to surprise on upside in the near term. Next generation platforms (defined as social "networking + mobile") are seen as driving unprecedented change in communications and commerce. In fact, mobile internet has already outpaced desktop internet adoption, with global mobile IP traffic expected to grow a staggering 66 times by 2013 and heavy mobile data users are expected to reach 1 billion in the same year. It seems quite reasonable to say that content and creative industries need to pay more attention to mobile than ever before.

3.2 Domestic Digital Media Technology

3.2.1 "The Digital & Connected Living Room"

Since the advent of the Internet, consumers have moved steadily away from the living room television which for many years, acted as the main hub connecting viewers to the rest of the world. However, this trend has been reversed lately with the popularity of high definition televisions, which allow consumers to connect many media devices to their televisions via the *High Definition Multimedia Interface* (HDMI). Access to the web and social networks on TV is now possible using devices such as Wii, Slinbox, Vudu, TiVo, Apple TV or even the digital cable *set-top box* (STB). This development has turned the television and living room into a media centre well beyond traditional TV viewing in the home.

The *Digital Living Network Alliance* (DLNA) certification is a new way to create networks between different devices such as PCs, consumer electronic products and mobile devices. It lets consumers easily acquire, view and manage an increasing variety of digital content from anywhere in or outside the home. This digital network gives consumers the freedom they want, while providing manufacturers with an opportunity to innovate and differentiate their products.

Setting a complex inter-connected network of media devices can be a challenging task for most consumers. The *Wireless High Definition Interface* (WHDI) aims to provide a wireless network which allows high definition video signals to pass between media devices without the need of complex and cumbersome wiring. It is adopted by major companies such as Hitachi, Motorola, Samsung, Sharp and Sony, making it a stronger contender to be the future wireless consumer electronics interface in the home.

At present, different network topologies/speeds are available depending on country and consumer choice. An *Asymmetric Digital Subscriber Line* (ADSL) with more than 6 Mb can guarantee access to certain content. But quality will be low and the receiver will need buffering to smooth out the effect of bandwidth fluctuations. Faster networks such as *Very high bit rate DSL* (VDSL) at 30-40 Mb and *Fibre To The Home* (FTTH) at up to 100 Mb could pave the way for higher resolution (e.g. HD) video reception in homes, and even make it possible to receive more than one unicast channel on the same connection through multiple set-top boxes.

To reach TVs via the Internet, receivers must have *Internet Protocol* (IP) connectivity. These receivers can be classified by environments:

Closed Environments

In closed environments all content, the header and the receivers are linked and controlled by the network operator. Decoders are identified here, for example, *Imagenio* to Spain. Each operator covers the cost of the receiver, so they implement their own IP stack. This is a pure streaming environment, that is, *Real Time Streaming Protocol* (RTSP), offering multicast channels, a pay-per-view library and interactive applications.

Semi-Open Environments

TV manufacturers hit the 2009 market with products offering new functionalities and network connectivity. Thanks to the DLNA protocol and device connectivity, they began integrating IP connectivity in TVs. The second step was to display the content available on the Internet. A common characteristic among manufacturers is to retain some control over the applications and content available, hence, the name “semi-open”, implying some control over the information published. They offer portals controlled by the manufacturers themselves and access is given to information or widgets, adapted for screen consumption. So we are referring to a vertical widget market for each TV, with prior manufacturer validation. In this direction, *Yahoo!* has announced that, thanks to agreements with Intel, it can offer an interactive platform for different TV brand names using the same CPU. There are many different codecs in the web market, and these TVs usually decide on a list of the most-frequently used in order to cover the biggest market share at optimal costs.

Open Environments

The term “open” refers to receivers that are not limited to an IPTV environment; therefore, the operator does not take on the cost of the receiver, nor do they need to be controlled by manufacturers. In open environments, we find hybrid broadcast receivers that use ethernet connectivity with an embedded navigator providing total freedom for internet connection. However, manufacturers expect webs to adapt their format to TVs to get more visitors and advertising (e.g., *YouTube XL*). Manufacturers that take this approach are Wyplay and Netgem, both are part of the *Open IPTV Forum*, and Amino will probably join them in the future.

3.2.2 Blu-Ray Adoption

In 2009, there had been much improvement on sales for Blu-ray hardware and contents, but with the economic situation and other factors, the format has not seen full adoption yet. Despite being an older technology, the DVD format currently still posts as a strong competition to Blu-ray. In the US, Blu-ray only accounts for 6% of the total home entertainment revenue¹. Part of the reasons would be cheaper DVD players, discs and movie titles compared to the Blu-ray counterparts. In addition, most modern DVD players now include video up-scaling as standard, providing improved picture quality to match the popular, large LCD TVs. In effect, the Blu-ray technology’s overwhelming picture quality advantage over

DVD has slightly lost its edge on averaged consumers. Another contributing factor might be consumer awareness, as a recent survey found that many people still do not realize a Blu-ray player could play DVDs too.

When compared to Internet, cable and satellite videos, the Blu-ray versions undoubtedly would have higher picture and sound quality, but many consumers find these download/on-demand services both convenient and a lot cheaper. In fact, charges from some US online rentals could be as low as \$1 per unit¹, making it difficult for discs in general to compete in this economic climate.

However, there have been some very positive results for Blu-ray too. Firstly, player prices have come down significantly, possibly due to the economic downturn. A few “entry level” players are now expected to sell for under \$100 in the US during Christmas 2009 and some Blu-ray titles could be under \$10 each¹. In addition, many (Profile 2.0) players now include internet connection (BD-Live compatible) to access online contents to supplement the movie on disc. *YouTube* streaming and DivX compatibility are becoming popular features which allow consumer to view free online clips and downloaded movies on television sets. TVs with integrated Blu-ray player are also on the horizon and standardization on 3D Blu-ray format is scheduled to finish in 2009, which could result in a big leap forward for digital home entertainment.

The growth in 2009 sales is also encouraging. As an indication, the Digital Entertainment Group in the US reported that total discs sales for the first three quarters was \$568 millions, representing 88% rise over the same period in 2008². The total number of playback devices (PlayStation and standalone players) sold in the US during first 9 months was 3.3 millions units, an increase of 13% over the same period in 2008².

In September 2009, Sony launched the smaller and more energy efficient *PS3 Slim* Console. This product should help to promote Blu-ray technology and potentially increase disc sales too, but it is also a serious competitor for standalone players because of its massive storage and additional usage as a high definition gaming console.

In the mean time, although the high definition disc format war has long ended, the delivery format war is still on-going and much work is still to be done for Blu-ray to dominate.

3.2.3 Games

Domestic games platforms currently fall into 3 categories – consoles, including handheld devices, PC and Mobile; however cloud computing is emerging as a potential delivery method, direct to the television. The *OnLive* service, an example of ‘cloud computing game platform’, was announced at the 2009 E3 games conference and is currently in Beta testing.

There are 30 million *XBOX 360s* now in homes around the world and Microsoft announced in November 09 that 10 million 360s had been sold in Europe, the Middle East and Africa with the balance largely owned in the USA. The 360 is 20 million behind the *Nintendo Wii* with 50 million machines sold to date and the *DS* handheld console has 100 million units in circulation. Sony’s *PlayStation 3* has sold more than 24 million units, while there are 53 million *PSPs* sold around the world. The *PSP Go*, the latest iteration of the PSP brand, was released in October. Like the *DSi*, it features enhanced connectivity with Wi-Fi and Bluetooth to drive downloadable content or interacting with other users, thus creating communities of players to tap into social networking trends.

Wii sales have fallen 45% when compared to the previous year, but sales of the 360 and *PS3* have increased, largely driven by price cuts in a difficult consumer market. Global sales of console games software across all platforms have dropped by more than 30% over the last year.

The launch of the *iPhone* has seen a resurgence of interest in mobile games. There are currently 100,000 ‘apps’ available on the *iPhone* store, of which roughly 10% are games. Game data tracking service Fade LLC reported earnings of \$250m from game sales over one year since launch in July 2008. 10 – 15% of the games hosted take 90% of the revenue, making it difficult for new titles, or those without a well known brand attached, to make an impact. The Google Android market is growing with 10,000 apps currently available and the new Microsoft Windows *Marketplace for Mobile* will be launched with *Windows Mobile 6.5* shortly.

Two key trends for consoles continued from 2008 into 2009. Further launches of games with peripherals and expansion of services utilising broadband connectivity. *XBOX 360* owners can now access their *Facebook* accounts, *Twitter* and Sky TV as well as downloading and streaming HD movies; The 360 is also expected to get a Blu-Ray player add in option; *Wii* owners can access the internet and the BBC

iPlayer direct from their console; *PS3* owners have full internet access as well as the option to download or stream HD movies and the *PS3* could be used for gaming or as a Blu-Ray player. Further developments in this area can be expected. The *Guitar Hero* and *Rock Band* franchises have expanded by offering new music packs with specific artists and new musical instruments, the *Metallica Guitar* for example for *Guitar Hero Metallica*. *DJ Hero* has just been released with a turntable as the game controller. Recent industry reports suggest this trend has run its course.

2010 will see further experimentation with more intuitive interfaces, which was started by the *WiiMote*. Microsoft's *Natal* depth sensing camera system, which promises controller free game play and interaction with the game through movement alone will be launched late in the year. Also coming in 2010 is Sony's Motion Controller and as the name suggests, it retains a control device but also uses motion sensing and a camera to interact with the game.

True stereoscopic 3D games will be a growing theme in 2010. All the major TV manufacturers are announcing plans for integrating 3D displays at the Consumer Electronics Show in January 2010. NVidia launched their 3D display system for PC games in 2009 and most of the blockbuster family movies planned for release from 2010 will be in 3D. Blitz Games Studios released world's first downloadable 3D game in August 2009 and expects this trend to gather pace, although widespread uptake is unlikely to take place until good auto stereoscopic displays (glasses free) become available.

3.2.4 User-Generated Content

User generated content (UGC) has more or less been encompassed by the umbrella of social networking in analysts eyes. Creating content and sharing it has become a major conduit for interpersonal communication whether it be by sharing photos on *Facebook* and *Flickr* or creating 3D chat avatars on *IMVU*. It is very interesting to note though that major changes in consumer technology seem to be coming in 3-5 year cycles with major industry players suffering major upsets in these paradigm shifts. The most obvious examples in recent years are the decline of *MySpace* as a result of the emergence of *Facebook* (in 2005 Rupert Murdoch's News Corp bought *MySpace* for \$580 million; in 2009 industry insiders declared the company 'pretty much worthless') and the emergence of Apple in the mobile internet market, making dramatic inroads into a space previously dominated by Nokia and Sony Erikson.

Given a flurry of industry trend reports led by Morgan Stanley, it seems quite apparent that user generated content and social networking is moving towards mobile internet devices. Japan may well be considered world leader in this area, at least as far as its domestic market is concerned with users now more likely to access the internet from cell phones than PCs. As of 2008, over 50% of middle school children and 31% of elementary school children had their own cell phones and many of them used phones to access the internet. In the west, the ability to post video and photos directly from *iPhones* to *YouTube* and *Facebook* has certainly been a factor in driving the success.

In the second decade of the 21st century, it seems clear that the media industry seems set for further paradigm shifts; mobile internet and location based content, both provided by advertisers and users, seem likely to be pre-eminent.

3.3 Professional Digital Media Technology

3.3.1 Introduction

It is more and more obvious that analogue technologies in the media market are vanishing. Many countries banned analogue transmission on terrestrial broadcasting and many satellite operators will do so in the near future. Although not all planned stations got into operation as scheduled, it is safe to say that analogue over-the-air is gone for the U.S. and many countries in Europe. With analogue transmission leaving behind, the industry can now better aim at more targeted improvements for implementations in capture, manipulation, distribution and presentation for the moving image.

Although currently the 4:3 footprint seems to be a format for eternity, a significant growth in HD for news, production and consumer applications can be also seen. It is very common to have image materials re-purposed simultaneously for different platforms and disseminated on different transmission media to another set of users. With this evolution, the lines of perception for image quality have become less definable and more dependent on the technology involved in the production chain.

The moving media industry continues to look for improvements at all stages of the content life cycle. As progress has been made in developing higher resolution imagers, recorders, and such, it is also being made in ancillary components that surround the overall media production chain.

3.3.2 Acquisition

Camera equipment and file/tape delivery of content result in a variety of physical formats and even greater variety of wrappers and compression formats arriving in a facility daily. Efficiencies in the workflow process can be achieved if a common versatile wrapper is used within the facility. These efficiencies are enhanced if a common compression codec can be maintained throughout as much of the facility as is possible. For SD workflows, 50-Mbit/sec I-frame MPEG in an MXF wrapper is becoming the de-facto intermediate format. For HD, there is currently no winner, but MXF is becoming the de-facto wrapper standard, with *QuickTime* and *AVI* following close behind. Codecs range from high bit rate, 10-bit JPEG2000 at the high end to low bit rate MPEG 4:2:0 at the low end. *AVCIntra*, *DNxHD*, *Pro-Res*, *I-frame MPEG-2*, *HuffYUV*, *CineForm*, *Long GoP H.264*, and *Long GoP MPEG-2* are all used in workflows around the world³.

Digital Camera development has reached the point where cameras are being built for use in theatrical motion picture production as opposed to primarily for TV. At the end of 2008, Sony Pictures Entertainment and subsidiaries was shooting about 20% of its pictures digitally.

As images get larger and sharper, the requirements for high resolution lenses get more definitive. For sports applications, focus is critical and auto-focusing has become a key element of the overall acceptance of a camera. New interesting technologies about this item appeared, as Canon placed its TTL-Secondary image registration Phase Detection technology in a new developed HD studio lens. The phase detection technology divides the incoming light into a pair of secondary image lenses focused on separate sensors, which then compare the two images to solve focus parameters for the primary image.

One trend seen in production is the merging of digital still photography and video onto a common platform. One example from Canon is the industry's first full frame, 21.1 megapixel (5616x3744) EOS digital SLR camera, with full HD 1920x1080 video capability at 30frames/sec. Panasonic's *Lumix DMC-GH1* with 12.1 megapixel imager is able to record both 1080p24 and 720p60 HD video. The new designed lens supports a multi-aspect ratio function, allowing photos to be taken using a 4x3, 3x2 or 16x9 aspect ratio with the same angle of view⁴.

Digital cinematographers, directors of photography and production professionals in all industries have discovered new uses of electronic cameras, such as production framing and checks as lighting, set of continuity, prop/dressing placement and the like. Others use them as conventional video recorders to gather HD images suitable for editorial purposes.

When capturing in 3D, additional requirements have to be fulfilled in the future. 3D viewfinders/monitors must allow directors of photography to properly correlate cameras and preview the results in anaglyph mode. 3D acquisition further demands new tool sets that can visually alert the operator of unacceptable disparities caused by too much out-of-the screen effect or background offset, depending on the size of the viewing screen used.

A number of companies are developing components allowing users to adapt pairs of the prominent manufacturer's cameras for shooting. However, an interesting approach for 3D capturing has been undertaken by SONY. The company announced the development of a single lens 3D camera technology that is capable of recording natural and smooth 3D images of even fast-moving subject matter such as sports, at 240fps. The technology combines a newly developed optical system for single lens 3D camera, which captures the left and right images simultaneously, together with existing high frame rate recording technology to realize 240fps 3D filming. Sony demonstrated a prototype model incorporating this technology at CEATEC Japan 2009. The introduction of a single lens system may resolve issues that occur as a result of having different optical characteristics for each eye. Furthermore, by using mirrors in place of shutters, incoming light can now be simultaneously separated into left and right images and recorded as it reaches the parallel light area of the relay lens. The separated left and right images are then processed and recorded with the respective left and right image sensors. As there is no difference in time between when the left and right eye images are captured, it is possible for natural and smooth 3D images to be captured, even for scenes involving rapid movement.

3.3.3 Storage

General

As with any other business, the media industry is paying increasingly more attention on gaining an edge by reducing operational risks. As a result, substantial technological advances in data storage to facilitate better data management and recovery can be seen. Enterprises are either investing in in-house storage solutions or are outsourcing their data backup needs, as they are looking to improve their ability to recover from any disaster at the earliest to ascertain business continuity. Within network storage, 'Storage Area Network' (SAN) is the most popular for its performance benefits, especially in mission critical applications. 'Network Attached Storage' (NAS) is the solution of choice among customers with specific file-serving requirements. Although 'Direct Attached Storage' (DAS) solutions are expected to compete head-on against the other solutions due to their low initial costs, network storage provides better returns in the long term.

Hard Disks (HD)

Enterprises struggle with the unabated growth of digital content. Therefore, maximizing storage rack-density and floor tile space, while simultaneously reducing energy consumption, are top concerns. For this reason hard disk manufactures further optimized the performances of current systems. The expansion in performance results in a better GB-per-Watt and cost-per-GB parameter.

For instance Hitachi Global Storage Technologies began to ship the industry's first enterprise-class, 7200 RPM 2TB devices with mean-time-between-failure (MTBF) of 1.2 million hours for use in demanding 24x7 *nearline* (or near-online) applications⁶. This new disk has double the capacity of the prior generation drive in the same 3.5-inch form factor.

As digital rights management (DRM) becomes more and more an issue among media providers, safe storage becomes an item of interest. For this reason some drives are available with a bulk data encryption (BDE) option. When enabled, the BDE implementation encrypts all data on the drive using a private security key as it is written to the disk, and then decrypts it with the key as it is retrieved, giving users an extreme level of data protection. Unlike software-based encryption solutions, hardware BDE implementation does not slow the system down. In addition, this also speeds up and simplifies the drive re-deployment and decommissioning process. By deleting the encryption key, the data is rendered unreadable, thereby eliminating the need for time-consuming, multi-pattern data overwrite.

Solid State Drives (SSDs)

It is expected that solid-state drives reach price parity with hard-disk drives within the next few years amid steep annual price declines in flash memory chips. Solid-state drives, which use flash memory chips as the storage medium, offer much better performance than hard-disk drives for some applications. Currently, opting for an SSD instead of a hard-disk drive will add anywhere between €60 and €400 to the cost of a laptop, depending on the capacity of the SSD⁵.

According to announcements of major flash manufacturers, reaching price parity with hard-disk drives is just a matter of time. Flash memory in the last five years has come down by 40, 50 and 60 percent per year. It is expected that Flash on a dollar-per-gigabyte basis will reach price parity, at some point, with hard disk drives in the next few years.

Due to the fundamental differences between the solid state technology and hard disk technology the break even barrier is very much depending on desired storage size per unit. A hard drive has a fixed cost of €30 or €40 for the spindle, the motors, the printed circuit board and the cables, so making faster or higher capacity hard drive would only adds little incremental cost to the drive. (The price for most laptop-class hard-disk drives on the market is between €40 and €70 at retail). SSDs also have a fixed cost for the printed circuit board, the case and the controller but these costs are significant lower. On the other hand costs scale linearly by scaling up the capacity of a SSD.

Tape

Most customers generally prefer the faster backup speed of disk-based backup and also the very much faster restore speed from hard drive arrays. However, where there is an enormous amount of data to be stored – Petabytes of the stuff - then bulk tape libraries still have a lower cost than the equivalent deduplicated drive array storage alternative.

There are several tape formats on the market but it seems that the LTO technology is the only technology with a defined upgrade path for the mass market. The *Linear Tape Open* (LTO) consortium has three technology-providing members: HP, IBM and Quantum. It defines tape formats against which the three members build drives and independent licensed manufacturers manufacture tape media and build drives. The current format is LTO-4, with a raw capacity of 800GB. The LTO development roadmap has LTO-5 as its next format and this is now licensed to manufacture. Its raw capacity is twice LTO-4's, 1.6TB, with a 2:1 compressed capacity of 3.2TB. LTO-3, the prior generation and format to LTO-4, had a capacity of 400GB raw and an 80MB/sec transfer speed. The transfer speed rose to 120MB/sec with LTO-4 and will rise again to 180MB/sec with LTO-5⁷.

There is a sixth generation of the LTO format on the roadmap, featuring a raw capacity of 3.2TB and 270MB/sec transfer speed. With the development of deduplicated and replicated high-capacity SATA hard drive storage arrays, where today's 2TB drives can hold 20TB of deduplicated backup data or more in virtual tape libraries or as straight filer-interface disk backup stores.

3.3.4 Post Production

It is apparent that software based post production is the norm and no longer an exception. The non linear editing process is now squarely set in the direction of an all file-based post production environment. Software applications that move metadata from set generated by cameras and other equipment into non linear editing platforms (e.g. *Final Cut pro*, *Avid*, *Bones Dailies*) are becoming a necessity. The Material Exchange Format (MXF) standard is firmly established in Broadcast Post-Production, mainly because of the success of the Sony *eVTR* and its registered document disclosure (RDD) to the Society of Motion Picture and Television Engineers (SMPTE)⁴. Workflows combining the *eVTR*, *Avid NLE* systems, and broadcast servers using MXF in coordination with Advanced Authoring Format (AAF) are now possible. Post Production requirements now emphasize the importance of features such as unwrapping and rewrapping, or the conversion of *QuickTime* compatible media to standard MXF in D10/IMX as built-in applications, plug-ins, or third-party application.

To overcome the format difficulties, IBM combined their MXF-server and General Parallel File System (GPFS) platform into a single harmonious working environment enabling the mixture of *Avid*, *Final Cut Pro* and *Adobe* software suites⁸. The GPFS incorporates high performance enterprise file management, allowing the user to move beyond simply adding storage to optimize data management.

3.3.5 Industry Trends

Digital Cinema

Digital cinema was growing in 2009. In the US more installations took place by mid-2009 than during the entire 2008. It is assumed that the number of digital screens worldwide has passed the 10,000 mark at mid-year. It can be seen that the growing digital footprint is scattered with only a few digital screens per complex, therefore distributors do not currently benefit from this technology. This year the top three circuits in the US announced their intention to install 4K projection technology, although having committed to 2K technology only a year ago. It is worth mentioning that there is currently no standard for 4K 3D projection, as SMPTE standards and the Digital Cinema Initiative (DCI) specification permit only 2K 3D. It seems that sights are set on 4K technology for capturing, scanning and screening.

3D Cinema

The evolution of stereoscopic 3D content continued in 2009. This year has been the turning point with the release of more than twelve movies featured in 3D and nearly twenty four 3D movies are scheduled for release in 2010.

Despite of several failed attempts in the past decades, it seems to be obvious that the 3D technology is here to stay this time. Technological progress was demonstrated from high end camera rigs to consumer displays. Directors of photography demand 3D cameras that will allow them to shoot as fast and efficiently as they are able to shoot in 2D. In addition 3D post-production tools are needed to meet deadlines and budget with no drawbacks in quality. This requires new cameras and production systems which are designed specifically for 3D to be available in the near future.

On the distribution side, big industry names such as SONY and Barco announced large supply contracts of digital projectors for theatres¹⁰. These companies are going to install more than 20,000 screens by 2012. The figures for new projection systems appear to indicate that the 3D cinema "egg-and-hen" deadlock has already been resolved.

3D Home

There is a tremendous amount of activity under way to bring the experience of stereoscopic technology to the home, on notebooks, monitors, TVs and home theatre projectors using several 3D technologies.

The following list is just an excerpt of work and promotion achieved at this time of writing.

- Every consumer TV display technology (LCD, plasma, DLP, RPTV) is moving toward a 3D-capable 120Hz refresh rate with prototypes shown at 240Hz.
- 3D technology will soon appear on small screen consumer products such as cell-phones / smart-phones, camera systems and picture frames (creating the 3D ecosystem requiring content, a delivery mechanism and a display system).
- Panasonic has officially submitted a proposal for 3D implementation on the Blu-ray format¹³. Their submission is based on the system which was first shown during CEATEC in 2008, and requires the use of special glasses which convert two separate 1080i60 video streams into one 1080p24 3D image.
- FUJIFILM Corporation announced a radical departure from current imaging systems with the development of a completely new, real image system. The new 3D image system features advanced image signal processing and micro-component technologies, and is so far able to demonstrate a camera, a viewing panel and a 3D printing system. This marks a complete break from previous attempts to introduce this technology.
- English satellite operator BSkyB is planning to open a 3D channel in 2010 offering movies, entertainment and sports programmes in the UK. The satellite broadcaster has stepped up its roll-out activities after a major boost in the numbers subscribers to its Sky+ HD set-top boxes, which are capable of broadcasting 3D services. To watch the 3D programmes, customers will require a new "3D-Ready" TV set, which are expected to be on sale in the UK next year, and use special 3D glasses.
- NVIDIA has sold more 3D glasses than expected in Europe¹¹. Their new stereoscopic 3D gaming technology promises to bring high quality 3D gaming to the PC. The 3D Vision technology utilizes active shutter glasses and a 120 Hz display (either 120 Hz LCD or 3D-Ready DLP TVs) to bring an immersive 3D effect to PC games. Using the depth buffer information stored in DirectX, the NVIDIA software is able to construct a stereo 3D image out of existing game content while the 120 Hz requirement gives each eye 60 frames of motion per second negating the physical detriments that were known to occur with previous 3D offerings.
- Minoru announced first stereoscopic webcam. This is a first step allowing users to generate their own 3D content¹².
- Google *YouTube* has already started working on providing support for 3D Videos by allowing users to view 3D videos right from their own PC with a 3D Player¹⁴.

3.4 Video & Audio Standards

This section provides an update on video, speech/audio and media production standards which could be relevant to the research and commercialization activities of the SALERO consortium.

3.4.1 Video Standards

Digital Cinema

The previous version of this report (D10.3.3 in 2008) explained in detail different standardisation activities in the field of "Digital Film" and "Digital Cinema" technologies. No fundamental changes in these activities can be seen over the past twelve months.

The *National Association of Theatre Owners* (NATO) released V2.0 of its Digital Cinema System Requirements document in Feb.2007 and submitted a revision V2.1 with minor changes in Dec. 2008¹⁵.

As theatre upgrades for the digital cinema are quite expensive and elaborate, the Digital Cinema Initiative (DCI) submitted a compliance test plan to ensure interoperability of products and services of all industry participants. With the publication of DCI's Digital Cinema System Specification (latest: Version 1.2, plus three errata publications), and the publication of related standards from SMPTE, ISO, and other

bodies, it is becoming possible to design and build D-cinema equipment that meets all DCI requirements.

3D Video

Sensio, the DVD forum which defines the standards for DVD announced new technology standards for 3D. The new standard – called Sensio 3D – will be supported by most of the players which are able to playback interlaced signals.

The Society of Cable Telecommunications Engineers (SCTE) has initiated a new project called “3D over Cable” that will focus on identifying necessary or desirable changes to existing SCTE standards including transport protocols, to facilitate the provision of 3D content by cable operators. A separate project to examine the delivery of 3D content over cable networks was approved by the Society’s Engineering Committee and assigned to SCTE’s Digital Video Subcommittee (DVS). As part of this effort, DVS will consider not only cable industry activities but also standards work being conducted in other organizations such as the Society of Motion Picture and Television Engineers (SMPTE) and the Consumer Electronics Association (CEA)¹⁶.

The Web3D Consortium is a non-profit organization that aims to reduce fragmentation in the industry with its X3D standard. This is a scalable and open software standard for defining and communicating real-time, interactive 3D content for visual effects and behavioural modelling. It can be used across hardware devices and in a broad range of applications including CAD, visual simulation, medical visualization, GIS, entertainment, educational, and multimedia presentations. X3D provides both the XML-encoding and the Scene Authoring Interface (SAI) to enable both web and non-web applications to incorporate real-time 3D data, presentations and controls into non-3D content. More information on this standard can be found at: <http://www.web3d.org/>.

Online Video

Online video standards continue to be based around the ubiquity of the Adobe Flash format. With the rise in popularity of video sharing sites such as YouTube, which displays all of its videos in the Flash format, the Flash plug-in for web browsers has become a default install for almost all web users, and thus the de facto default for all online video. This rise in popularity of Flash video sites has seen an equal rise in popularity of standalone plug-ins that enables users to embed Flash video in their own websites. Many commercial Flash players are available, but there are also open source alternatives such as *FlowPlayer* (see <http://flowplayer.org/> for more details).

HDMI – Version 1.4

A new version of HDMI was released in May. HDMI1.4 increases the maximum resolution to 4K x 2K (3840x2160p at 24Hz/25Hz/30Hz and 4096x2160p at 24Hz, which is a resolution used in digital theatres).

Major differences to HDMI 1.3 are

- Updated list of CEC commands
- 100 Mb/s ethernet connection between the two HDMI connected devices
- Introduction of an audio return channel
- 3D over HDMI
- New micro HDMI connector
- Expanded support for colour spaces
- Automotive connection system

TV Standards

The transition away from videotape to a more IT-oriented workflow for the acquisition and post production of TV and motion picture content has created the chance to make things simpler when it comes to synchronization and time labelling. The European Broadcast Union (EBU) and the Society of Motion Picture and Television Engineers (SMPTE) formed a joint task force to study the issue and a report was published in Q4 of 2009. The report summarizes nearly two years of industry research and user inputs to provide a comprehensive set of recommendations for simplifying and codifying synchronization systems and time-related labelling in the digital era. It features input from broadcast,

post-production, movie studio, and cable professionals, as well as broadcast and network equipment manufacturers. This report will set the requirements and framework for SMPTE to develop a revised standard, which will get under way in 2010¹⁷.

Display Technology

SMPTE recently completed a study needed to ensure that content creators remain in control of the appearance of their output. This work was judged necessary because of the imminent disappearance of the universal CRT display technology and its replacement with an array of multiple, disparate flat panel or competing fixed pixel matrix technologies. The upcoming standardization work will include the specifications needed to achieve interoperability between the content being judged and the monitors intended to perform the reference or quality control role, as well as definitions of the necessary supportive measurement and calibration procedures. One goal of the standardization activity is to fulfil the requirement that the standards and recommended practices to be issued are independent from the technology used⁴.

3.4.2 Speech and Audio Standards

Speech Standards

Speech analysis and annotation formats are often ad-hoc in terms of their application, notably in relation to metadata and content retrieval. The only cohesive attempt at corpus metadata standardisation is the ISLE Metadata Initiative (IMDI). Although not a comprehensive (or universally adopted) standard, IMDI represents the only current standard for speech corpus metadata available and is implemented by the DIT Emotional Speech Corpora (Salero documents D6.2.1 and D6.2.2). For acoustic and linguistic analysis, a broad distinction can be drawn between analysis of real speech and the processing of speech content for synthesis. In the case of the former, MPEG-7 provides basic spectral and temporal descriptor support but general media formats such as SMIL (as used in SALERO) allow for more verbose descriptions of attributes such as voice quality. In addition, the adoption of SMIL allows sequencing and synchronisation with visual content to be performed in a single file as demonstrated by the online character animation tools (D6.3.1). In the case of speech synthesis, SSML is a widely used W3C standard for the delivery of processed speech and is used by speech synthesis work within SALERO (see D6.4.1, D6.4.2, D6.4.3). SSML is also fully compatible with SMIL, and thus interoperability can be maintained with both real and synthesised content.

Audio Standards

Many audio standards such as those in MPEG1, MPEG2, DTS and Dolby have been around for many years and are used widely for audio compression in many discs, broadcast and Internet applications. The most recent version of MPEG-4 Part 3 (ISO/IEC 14496-3:2009), published this year, aims at interactive digital audio applications. The standard integrates different types of audio coding: natural sound with synthetic sound, low bit-rate delivery with high-quality delivery and lossless coding, speech with music, complex soundtracks with simple ones, and traditional content with interactive and virtual-reality content.

The Digital Living Network Alliance (DLNA) is set up in 2003 by a cross-industry group of consumer electronic companies. The goal of DLNA is to enable cross-industry convergence by establishing a platform of interoperability based on open and established industry standards. DLNA Certified® devices include televisions, cameras, printers, PCs, set-top boxes, cell phones and games consoles, etc. The audio part includes defining mandatory/optional audio codecs to be supported by certified devices. There are now over 5500 DLNA Certified® devices worldwide, offering consumers a greater variety of options for storing and sharing digital media.

Digital Entertainment Contents Ecosystem (DECE) is another standard proposed by a group of content and device companies (many companies are also members of the DLNA). The goal of this global standard is to allow consumer to buy the digital content once and play anywhere, anytime. It has the support of over 40 leading digital entertainment and technology companies including Warner, Sony, Intel, Microsoft, etc. One challenge for DECE: Apple is not part of the group and the success of iPhone (which uses its own digital rights management) could create problem for the adoption of DECE. The audio standardization part of DECE covers mandatory and optional audio codecs to be supported for standard/high definition players and portable devices.

Some of the main audio related IPTV standardization activities include the Advanced IPTV Terminal (AIT) standard and those carried out by the Open IPTV Forum (OIPF). The AIT standard, by MPEG,

aims to develop an *ecosystem for media value chains and networks*. A list of 30 “basic” services have been identified and are being defined. These services include “protocols” (payload formats) to enable users to call these services, “Application Programming Interfaces” to access services, and “bindings” to specific programming languages.

The Open IPTV Forum was created in March 2007, to provide an IPTV solution which enables a “plug and play” experience for the end-users and fills an industry gap making it independent from the technology behind it. For audio, it specifies mandatory and optional audio codes to be used in the delivery chain. Release 1, “Technical Specification”, was published in early 2009, followed by Release 2, “Architecture Specification”, in September.

3.4.3 Production Standards

The main production standards for broadcast media have not changed since their introduction several years ago. Analogue standards such as MUSE, NTSC, PAL and SECAM have all been widely used for decades. In recent years digital standards have been developed to assist the rise in use of digital technology:

- ATSC is the acronym for *Advanced Television Systems Committee* for Digital Television. Developed in 1993 by the Grand Alliance, a consortium created to develop American digital television and HDTV specification. It is replacing the analogue NTSC television system.
- DMB: The *Digital Media Broadcasting* standard was developed by South Korea for mobile multimedia broadcasting. This standard has recently been used by other countries such as Norway, Germany, France, etc.
- DVB: The *European Telecommunications Standards Institute*, the *European Committee for Electro technical Standardization* and the *European Broadcasting Union*, are in charge of maintaining what is, along with ATSC, considered the most important standard for *Digital Video Broadcasting*.
- ISDB: The digital television standard in Japan, and many south-American countries, the *Integrated Services Digital Broadcasting* replaced the *MUSE* analogue high definition television system.

In the post-production world there is a marked lack of file-format standards that are widely accepted. This is undoubtedly due the competitive nature of the industry, with each company pushing their own file formats and standards. In video production world this is especially the case as the switch to full digital technology was made less than a decade ago, and even now interviews with successful post-production houses such as Lapospo in Barcelona (www.lapospo.com) reveal that 10% of their work is still carried out on 35mm film.

Probably the widest accepted image format standard in the production industry is the *Digital Picture Exchange* (DPX) format. It was commonly used to represent the density of each colour channel of a scanned negative in a 10-bit uncompressed log format, although it also supports other video formats. The format was originally derived from the output of the Kodak Cineon ‘FIDO’ film scanner, and has now been published by SMPTE as ANSI/SMPTE 268M-2003.

A common output file format of many digital video cameras is the *JPEG-2000* image compression standard and coding system. Created by the Joint Photographic Experts Group in an attempt to supersede the original (and very widely used) JPEG format, it uses wavelet to compress each image and is highly flexible in how it may be decoded, permitting the user-application to sample the code-stream in different areas in order to rapidly obtain different representations of the image. It has been published as ISO/IEC standard 15444. JPEG-2000 also uses the EXIF standard to store metadata within each file. Although widely used in production, it has still not overhauled the popularity of its older brother, JPEG, for other applications such as consumer digital cameras and on the World Wide Web.

Finally the Tagged Image TIFF image format is still used widely in the production industry. While technically not an open standard (it is under the copyright of Adobe systems), its specification has been published as TIFF/EP (ISO 12234-2) and TIFF/IT (ISO 12639), as minor extensions to the original TIFF 6.0 specification published in 1992. TIFF is widely used in several fields of image production, including the consumer field, as is used as a direct output format of several high end production video cameras.

There are no globally accepted standards for videogames. Currently, for console games, routes to market are controlled by proprietary platforms and they each set their own technical standards. While the PC platform is an open one, games must be written to run on the chosen PC specification and operating system.

4 Digital Media Production Trends

4.1 Audiovisual Market Production Trends: Cross Media and Movies

This chapter deals with trends in the movies and cross media field, whilst production trends for television are described in chapter 4.3. In last years report²⁸ it was indicated that a continued growing trend in digitalization can be identified in the audiovisual market, allowing for more alternative content to be produced and leading to an increase in the use of Cross Media production formats, in which the same content or theme is expressed through multiple platforms. The producers' media field has become increasingly wide. Cross-platform thinking has become the norm within production companies.

4.1.1 Movies and 3D

In previous years' reports it has been shown that 3D – in it's various technological and artistic expressions – is gaining popularity on many media platforms. Cross-platform and cross-media thinking together with 3D technological innovations is subsequently affecting also storytelling. For instance, the major broadcaster HBO is creating demos of new formats, such as the *HBO Imagine* demo, featuring a multipoint-of-view show *The Affair*, which is presented on-line using the "HBO Cube" application.²⁹ Cross Media and interactive storytelling is also reaching edutainment for children; Mobile Art Lab in Japan, to give an example in this area, is developing cross media by combining a smart phone (Apple's iPhone) and printed media (children's picture book). The resulting project is the *Phone Book*, where the phone is inserted into the children's book, providing interactive audiovisual content within the book.³⁰

The popularity of 3D movies is on the rise, as predicted and regardless of the recession. Considering film audiences in general, for example in Finland the economic downturn has interestingly enough increased box office revenue by attracting people to visit cinemas as a "break from reality".³¹ 3D release technologies and digital movie theatre development has continued taking major steps forward. A wide variety of 3D technologies were displayed at the IBC 2009 Exhibition in Amsterdam this autumn, showing that much investment is directed towards the development of 3D, and that new technologies and standards for filming and projecting 3D movies are competing for the markets.³²

Several 3D movies have been released within the past 12 months. The film *My Bloody Valentine* by Lionsgate that premiered in January 2009 was projected in *Real D 3D* and released to 1033 3D screens, breaking the record of releases for this format. In addition it was released to 1501 regular screens, which shows that 2D is still complementing 3D on the film market. Disney seems to be a big player in the 3D market, and as expected, Disney Digital 3 D technology was used in the Walt Disney Pictures release *Bolt* (November 2008) and Pixar's first 3D feature film *Up*, released by Walt Disney Pictures and Disney Digital 3-D (May 2009). 3D technologies competing with the before mentioned are e.g. the stereoscopic InTru3D technology, developed by Intel Corporation in collaboration with DreamWorks Animation in 2008, as well as some technologies which still require the audience to wear specifically polarized glasses, like the Real 3D Cinema and IMAX 3D theatres that screened the movie *Monsters vs. Aliens*, released in March 2009. 3D projections are also combined with other immersive technologies that activate the viewers' sensory organs beyond the audiovisual; systems such as D-BOX simulating physical motion (in the movie or home theatre seat) and offering tangible sensations are already on the market. Movies supporting the D-BOX motion simulation system have been released at a number of cinemas in USA and Canada: *Fast and Furious* premiered in April 2009 and *The Final Destination* was released in Real 3 D and D-BOX in August 2009. The advertising and marketing sectors are picking up on the trend; a commercial screened in InTrue3D was seen during the February 2009 Superbowl. 3D is also entering the field of Cross Media installation art; in May 2009 the 3D stereoscopic film installation *Radio Mania: An Abandoned Work*, was commissioned by the British Film Institute, complete with 3D Ambisonic sound.³³

The use of 3D projections in installations and other artwork is within the reach of an increasing amount of artists and designers as affordable, professional level video projectors for High Definition 3D video and real-time images has become available on the market. The boundaries between movies, cross media entertainment and cross media art are fading, with social media being one of the major bridge-builders.

4.1.2 Cross-Media and Social Media Production Trends

As social media has gained popularity, users have become more accustomed to a variety of media forms and formats. End users continue increasingly to re-mix and re-use media content by using social media applications across different platforms (e.g. from web to mobile devices). Personalized media collections created and shared in this way can also be considered “everyday” cross-media. Programmes such as the UBI (Urban Interactions) Research Program, coordinated by the MediaTeamOulu at University of Oulu are studying urban computing, i.e. the interaction of urban space, computing technology and people. UBI aims to build new ubi infrastructure with large displays, sensor networks and software related to them, in central Oulu, bringing modern technology closer to the everyday users.³⁴ Developments such as these enable cross media to become a part of our everyday life also in urban spaces, and not just within our home entertainment systems, on our personal devices or in custom-made venues such as digital movie theatres.

Cross Media productions are increasingly utilizing open source software and shared knowledge of new tools. Also traditional television broadcasters have picked up on these developments and are implementing their own cross-media products and services. One example of new services developed by broadcasters is the Finnish Broadcasting Company, having tagged all their digitized programmes on their YLE Areena on-line site, so that they can be linked to social media applications (Facebook, Twitter etc.). TAIK is also currently collaborating with YLE in creating an on-line and mobile social media application concept for cultural services³⁵. Radiotjänst, the Swedish body responsible for collecting TV-licences, has recently released an on-line commercial, where you can upload a facial photograph of yourself onto the site <http://tackfilm.se>. It is then automatically processed and embedded into the commercial, as if you'd feature in the story yourself. Once the film is ready, you can forward it to your friends via Twitter, Facebook, or simply by e-mail or using the embed code. The Radiotjänst commercial works when using a 2D photograph, but similar 3D demos are to be expected to feature in the industry in future years.

Social (Cross) Media is also enabling citizens to participate in a more open societal and political discourse. Public access community IP or broadband television channels such as M2Hz (<http://www.m2hz.net/>) provide a platform for communities to create and broadcast their own content, while enabling transparent and open communication channels for issues and stories on multiculturalism, urban culture and events, experimental art, expressions of local life styles and culture – through audiovisual content produced by the viewers themselves.

4.1.3 Cross Media Interfaces and Cross Media Research

Cross Media research – for instance at research units at MIT, UPF and TAIK, to name a few – is showing signs of a trend of interest in touch-based, tangible, gestural and physical interfaces.³⁶ The translation from physical to non-material artwork via human-computer interaction and the interface has been explored by TAIK's Teemu Korpilampi in his MA New Media thesis work³⁷. TAIK's 3rd Experimental Production Turing Machine also explored user interactions and developed a touch-based interface for a physical installation. This 'Alan01' installation provided the audience with interaction possibilities – and thus an ability to affect the unfolding of the installation's media contents – via the touch screen interface, which was linked to a symbol database. Content-based search and retrieval and potential uses in cross media productions is described in further detail in the full paper *Concept, Content and the Convict* (available at the SALERO website www.salero.eu), co-authored by TAIK and UG. In the conclusions of the paper content-based image retrieval was found to show great promise as a tool for creating sophisticated non-textual installation and artwork interfaces. The paper was presented at the ACM International Conference on Multimedia 2009, Interactive Art Track on 20 October in Beijing, China. Nearly all the other presentations at the conference Interactive Art Track also utilised Cross Media and/or physical interfaces: http://staff.science.uva.nl/~nack/ACM_MM_IAP_09/papers.html [the site last visited Nov 26, 2009].

The discourse at ACMM 2009 suggested that further research in the area of projecting media content “beyond screens” should be encouraged.³⁸ Demonstrations of beyond screens cross media art is presented already by artist working on “urban screens”, in connection with events such as the forthcoming Media Facades Festival Europe 2010, co-produced by the Finnish association for new media culture, *m-cult*.³⁹

Research demonstrators created within IST co-funded projects are also being presented and tested by users in urban space. The IPCity project (FP-2004-IST-4-27571) which investigates Mixed Reality Interaction and Presence in Urban Environments has supported the Ubiquitous Interaction group at HIIT

in the development of *CityWall*, a large public multi-touch display in central Helsinki, which allows multiple users to interact with the content. The 3D touch interface is used by passers-by in the city, inviting them to view photos uploaded via Flickr by users in other cities, and allowing them also to upload their own images.⁴⁰

Planned cross media research at the research group Crucible Studio within TAIK⁴¹ aims to take an analytical eye on the existing visual/gesture communication in human created Art, Entertainment and Education, and describe them as "externalisations of mind" in the fashion described in PhD Pia Tikka's doctoral dissertation *Enactive Cinema: Simulatorium Eisensteinense*.⁴²

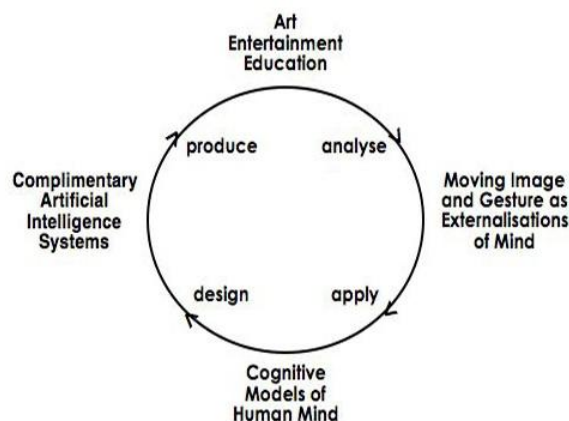


Figure 1: Graph representing the proposed process of future cross-media research at TAIK

The research would aim to frame - for purposes of mediation between human-machine intelligence - cognitive models or ontologies that can inform about the development of AI (artificial intelligent) that aids human interaction within a visual system, so that only gestural or visual communication is necessary. The research would develop applications and sample productions from **art**, like interactive theatre performances and museum installations, to **entertainment**, like toys, online story-worlds and games, and **education**, for example, for special needs children, art & design and complex information visualization. Also forthcoming PhD research within the research group focuses on the human-computer interaction through the development of various devices, interfaces and environments for key individuals with particular profound physical and mental disabilities chosen for the study; specifically on the creation and analysis of playful, participatory artworks and social environments that exploit both physical computing methods (such as embedded computers, complex systems, ubiquitous computing) and telematic systems (e.g. virtual environments, networked performance, telepresence).⁴³

The research group's special interest lies in the analysis in creating experimental productions/demonstrators for the AI system(s) to be developed. These could include virtual worlds as well as truly "Total Theatre" (of Richard Wagner, Gropius and Bauhaus) online and in physical space. It is also possible to continue within the themes explored in TAIK's 3rd EP Turing Machine, especially as the Alan

Turing centenary celebration year is forthcoming in 2012.

4.2 Games

The trend for high fidelity and realism, highlighted in previous reports continues. '*Call of Duty, Modern Warfare 2*', a mature rated realistic combat based game, became the highest grossing launch in entertainment history in November 09, selling 4.7 million units in the first 24 hours of release and grossing \$310m. The game has been controversial because it allows players to act as an undercover agent with a terrorist group, killing innocent people; for the first time warning screens are displayed before levels where morally ambiguous game play is offered and the player is given the opportunity to skip that level. This may be a pointer for future games where content which has been previously unacceptable is included, but with clear warnings about its nature.

A continuing strong trend is games for self improvement. *Will Fit Plus* and *The Biggest Loser*, (linked to the TV show) are just two of the current games in this genre that largely sit on the family friendly platforms such as *Wii* or *DSi*.

Character customisation also continues to be offered as a way of enhancing player experience by closer identification with the lead character. *'Fallout 3'* and *'Karaoke Revolution'* are two examples at opposite ends of the game content spectrum that both contain character customisation tools. *Fallout 3* is a mature role playing game set in post nuclear holocaust story whereas *Karaoke Revolution* is a singing game for the whole family.

Another parallel approach of importing real player likenesses via jpegs or web cam to 'paste' onto a head mesh is becoming increasingly successful. *Game Face* system from Electronic Arts (EA), takes front and side views of a player to create a 3D likeness which can be customised and saved it as an avatar that can be imported into any EA game. The first game using the system is the football game *'FIFA 10'*. The player's in game face has limited animations driven by player actions. Both approaches demonstrate the importance of procedural generation for today's games industry as a way of engaging players as well as promoting efficient production pipelines.

High production values continue to be important for driving sales at the top end of the market. UK based specialist post production house, The Foundry, produced the cut scene videos for *DJ Hero* and racing games like *Forza Motorsport 3* are promoted on their glamorous graphics as well as adrenaline pumping driving action.

However, in the current economic climate, the highest budgets are only being assigned to sequels or games linked to high profile film or TV licences. The industry is seeing some parallels with the film industry – big budget blockbusters and smaller budget indie productions. The continued growth of digital distribution direct to consumer via console, PC and mobile is supporting the development of agile smaller developers, whose low overheads can be met by the lower returns earned from these markets over a longer period than the boxed games recoupment cycle. High profile boxed products are enhancing their shelf life by including online multiplayer modes and access to additional downloadable content. In some games, there are opportunities to create new game modes and then upload these to the multiplayer server to share with friends.

Micro-transactions for virtual good in web based games linked to social networks, which have been well established in Asian markets, are now a driver for western markets too, with an estimated 30,000 social networking games now available. Large games publishers are watching these developments carefully for opportunities. In November 09, EA cut 1,500 jobs (about 17% of its work force) in the same week that it acquired UK based online social network games company Playfish for \$400m.

The rapid rise of virtual worlds, noted in 2008, has slowed with fewer worlds launched or planned to launch. One of the largest investors in *Second Life*, the Australian company Telstra, has recently announced it would be withdrawing from the virtual world and closing its *BigPond* sim, which used a membership subscription business model. The company told *BigPond* residents: "Our focus is moving towards entertainment options that are gaining in popularity on computers, mobile phones, TVs and game consoles."

2010 is likely to see an increasing crossover between console games, web and mobile games and social networks, with players connected to their game experience and their friends through any device, anywhere and at any time. Intelligent content creation, that enables re-use and automatic adaptation of assets to different delivery methods and platforms, will be increasingly important, as budgets continue to come under pressure in the continuing difficult economic climate.

4.3 Television

Although television is still the most preferred entertainment platform in Europe, new platforms (or new screens) are becoming a reality.

It appears that the future of this medium implies TV connectivity to the Internet, with broadcast/broadband hybrid customized services catering for different tastes. Viewers will be able to choose what they want to watch and when to watch it too. 3D television is already a reality, although it is still too soon to predict if it will be adopted on a massive scale or not. Given this future scenario, both telecommunication companies and TV manufacturers will play a new role as content producers.

IP-based Services

Television transmission over the Internet can be done in two ways: through an operator-controlled network (IPTV) or through the open Internet (Internet TV). Although direct connection to the Internet is not common at present, concepts such as IPTV or Internet TV are quickly becoming more and more

prevalent as of late and there are indications that it may become the technology that changes the future of television by offering viewers services and highly attractive content.

IP-based services offered on television:

- Live broadcast.
- Video on demand (VoD): this will probably be the highlight of this new TV culture.
- Time-shifted programming.
- Electronic program guides (EPG).
- Social TV - Content sharing, participation and communication with other users.
- Customized content and advertising.

Content that can be offered on IP-based television:

- TV channels
- A la carte content: Users can choose what they want to watch and when.
- Enhanced TV (additional content): Information complementing what the viewer is watching. For example, while watching a football game viewers can access related statistics or while watching a film they can access an interview with the film's director.
- Related content: Similar content which the viewer might find interesting. For example, while watching a documentary (e.g., on Spain's political transition), another documentary of related content is suggested (in this case, on the first democratic elections held after Franco's death).
- Personal content: Data stored on other home devices, i.e., PCs, cell phones, etc.
- Customized content: Content adapted to user preferences and profiles.
- Widgets ("Capsulized" content): Widgets are small applications with data that is updated from the web. They are containers that can offer all types of content: weather forecasts, stock reports, sports, photos, videos, etc.

Worldwide IPTV growth is accelerating, with the Informa Telecoms & Media group predicting 38.4 million homes by 2012; IMS Research group predicting over 52 million; ABI Research predicting up to 70 million homes by 2012 and about 90 million by 2013. According to ABI, Western Europe will likely continue as leader in the global IPTV market until 2011/2013 and then slowly being overtaken by emerging markets in the Far East, Eastern Europe and Russia, which are all growing in a very rapid pace.

In the United States the IPTV services offered by Verizon and AT&T are growing faster than expected. The Management Research Group expects the number of subscribers in the US to grow from 13.5 million in 2007 to 72.6 million in 2011, an annual market growth rate of 40%. Their profits for these years are estimated at between 2,100 and 2,500 million dollars (or 1,428 and 1,700 million Euros).

IP-based television is opening new business opportunities for content producers and broadcasters. Today, the main business models for IPTV services are '*Triple Play*' (high-speed Internet connection, IPTV and VoIP) and '*Quad Play*' (*Triple Play*, plus mobile telephones), offering consumers bundled services from a single provider. IPTV and Internet TV scenarios are paving the way for many other business models such as pay content (e.g. subscriptions, renting, pay-per-view) and free content with revenue from advertising, directed advertising, personalized offers, etc.

Hybrid Broadcast/Broadband Services

It is obvious that television is in the first stages of a transition to hybrid broadcast-broadband services. There are several research projects underway aimed at making this new scenario as standardized and open as possible. One of these initiatives is Hybrid Broadcast/Broadband (HBB) which covers devices and services to offer both broadcast and internet content. Broadcasters in France and Germany are promoting their own version which is based on open standards (HTML applications, JavaScript and CSS). They are working with the European Broadcasting Union's WMT (Web Media Technologies) group to come up with a first draft of specifications to present to the *European Telecommunications Standards Institute* (ETSI). In addition to broadcasters in Germany and France, IRT, Philips, Orange, Alcatel, Netgem, OpenTV and Astra are also participating in this initiative. The technical demonstration

that they have available shows a DVB-T television device overlaying content downloaded from the Internet onto the broadcast. This new version of TV, also known as *RichTV*, will allow broadcasters to modernize catch-up TV services such as ARD/ZDF's *Mediathek*, and the BBC *iPlayer*, in addition to a new HTML-based environment to generate teletext.

Users will have access to different interactive services based on the web and provided by a broadcaster. These services include news and subtitles in different languages, plus information related to the program being broadcast and weather forecasts. The additional information will be superimposed over the TV program.

Costs are expected to be similar to those of TV today given that there will be no licenses and broadcasters will not be subjected to a canon. Everything will be standardized by the ETSI to avoid market fragmentation and to allow for economy of scale.

Main broadcast features:

- Transmission of TV and radio services
- Applications related to the broadcast using DVB signalling
- Transport and synchronization of applications and related data (a teletext web could be displayed in HTML format even with receivers that do not have broadband connection)

Main broadband features:

- Video services on demand.
- Transport of associated applications and data that complement broadcast.
- Server-customer information exchange.

The applications can be presented thanks to a browser allowing the following:

- It must be open and not based on one authority control.
- Services and contents from independent suppliers accessible on the same terminal.
- Terminal's standard functionalities available to all applications.
- Services and content must be able to be subjected to viewer protection regulations.

The 3D Experience from Home

It appears that once HD television is consolidated, the next step will be 3D television. Manufacturers have their hopes set on this innovative technology and they think it will be less expensive than all the technological development needed to provide high definition content. Technological giants such as Sony, Hitachi, Panasonic and Sharp are working on the 3D system and they already have some prototypes that could be put on the market when there is more content available in this format.

Time will tell whether mass consumption of living room technology will take off or not, but there is no doubt that 3D TV is now a reality. In late 2007, Japan became the first country to create a 3D channel. The operator was Nippon BS Broadcasting through the Broadcast Satellite (BS) HD Channel. Today Sky wants to continue in that direction and it has announced that in 2010 it will offer a 3D channel with films, entertainment and sports. The British satellite TV platform has successfully tested recording programs in 3D – boxing, rugby matches, a program from "*Gladiators*", etc. The channel will be transmitted over their existing infrastructure in their high definition platform. Therefore, viewers with SKY+HD set-top boxes can watch it. But to enjoy the 3D experience, a 3D Ready TV will be needed. These are expected to go on sale in the UK in 2010.

But another of the future challenges of 3D technology is live re-transmissions. To date, all live 3D productions and re-transmissions have been in the testing stages and there is not yet a standard. In 2008 the BBC was the first network to offer a live rugby match – Six Nations Rugby between England and Scotland – which was broadcast in a London cinema. Mobile units and cameras will need technical modifications for 3D live retransmissions. One of the big advantages of 3D content is that it draws viewers into the action and gives them a more intense experience. This means that film directing will also need to be re-examined to come up with a more attractive visual language.

As we have seen, it seems that the technology is ready, but experts do not agree as to whether 3D will eventually be adopted in homes. The most optimistic estimates claim that by 2013, 46 million TV sets

that allow HD and 3D home viewing will have been sold. On the other hand, in a survey done by Sky among 427 European broadcasters, only 35% include 3D production among their future goals.

Holographic television may sound like something out of science fiction (such as those in the *Star Wars* movies), but during the last U.S. elections, the first hologram was seen on live TV. On election night in the U.S., CNN New York connected with one of their reporters on location in Chicago. To everyone's surprise, the reporter turned up in 3D in their New York studio. To create this holographic image, the journalist was reporting from a special tent set up in Chicago. There were 35 high definition cameras set up in the tent, filming her from 35 different angles. Right now, it is too early to talk about holographic TV as a technology that will eventually be adopted, but it is obvious that it can open new doors for television in the future.

Personalized Television

Although personalized TV in the web environment is already starting to take off, the next step will be personalized television on living room TVs.

The Internet now offers a number of services offering personalized TV consumption. Most of these services base their offer on the possibility of zapping from channel to channel (some are exclusive to the Internet, while others are the regular TV broadcasts). The possibility to access Video-on-Demand can also be added to this zapping service. The videos are categorized and users can choose which one they want to watch. This offer, which is not personalized, can also provide services to registered users with varying degrees of personalization. One of the most interesting personalized TV web services is ADNStream. This web offers what they call "*Mutant TV*". Users indicate their preferences when they register. This, together with their votes for the videos they watch, provides the information needed by the system to offer a personalized selection of channels and Videos-on-Demand to registered users.

As we have seen in reference to IP-based TV, televisions are also starting to be connected to the Internet. So all the possibilities for personalization can be applied to living room TVs as well.

To achieve personalized television, there is also another technological tendency that is gaining momentum: the possibility of identifying TV viewers. More and more projects are in the works to provide broadcasters and telecommunication companies with information on the users watching TV. An example can be found in the last agreement between NDS, one of the most important companies in technological solutions for digital television, and TNS Media Research, leader in measuring TV, radio, and Internet audiences. Both companies offer an audience measurement solution based on a pay TV decoder that identifies the individuals in each home.

The BBC has also been researching and experimenting in this direction for years and it has several projects underway aimed at identifying the users who are watching their channels through decoders.

Some research projects identify users by their cell phone numbers. Through wireless technology the decoder identifies which SIM cards are in front of the TV and thus, who the viewers are at that moment.

In addition, commercial interests in identifying viewers (since apart from content, this can also allow for personalized advertising) are likely to provide the financing needed in the years to come to develop technological standards permitting user identification, which is key in a personalized service.

In short, existing technology can already offer personalized content and advertising via the Internet. Viewer identification already exists, and so does the possibility of sending broadband content over an IP connection. Therefore, the technological channels needed to offer personalized TV exist, offering a customized programming schedule to meet viewers' tastes, for example, or recommending contents on demand according to viewer preference or audiovisual consumption habits.

New Agents and New Roles in the Audiovisual Market

As things stand, the technologies we have looked at have had little impact on the market. Even though more TVs and decoders that can connect to the Internet are being sold, renewing the TV set park is not a fast process and several years will probably be needed before the market impact is high enough to convince content producers to place their bets on these services.

In this new scenario, telecommunication companies and TV manufacturers can play a new and decisive role as content producers. In fact, many of them have already begun offering closed content packets, also called "*walled garden*".

Several reports have even predicted strong growth for telecommunication companies and TV manufacturers as content suppliers in this new TV via IP scenario, which will hurt broadcasters who are not well positioned in the television industry's technological transition stage.

The following headings in bold, within Section 4.3 are trends identified by the *International Institute for Television Leadership* in 2008 and could probably be valid for the coming years:

Multi-Channel Universe

TV is still Europe's preferred forum of media entertainment. In fact there has been significant fall-off in the movie-going activity for TV, gaming or Internet entertainment. Distributors of broadcaster signals and other offerings such as VoD and Sports channels are capturing most market share.

European markets are already starting to look at both production and distribution, based upon screen size/location, and are talking about the concept of distance from screen to viewer:

- 100 ft - Public Screen (outdoor/sports complex/bars, etc.) - "Event" content & HD format
- 10 ft - Home Entertainment & Engagement in the experience - HD programming
- 1 ft - Computer/Mobile entertainment - short programming cycles and options HD not required but multiple formats of carriers and end-consumer technology are necessitating multi-format productions.

The content and the writing required for each of these programming environments is quite different and many are developing stories or entertainment streams in parallel by different creative groups for the different environments right from the start.

Language and Geographic Boundaries are Dropping

With the growth of the EU, and the increased travel across EU and outside EU to Asia and Americas, the younger demographics are increasingly comfortable and capable in the English language. Most use it day to day in their work, on computers, and increasingly in their entertainment. In addition, many speak several languages such as German, French, Spanish, English, etc.

As a result the traditional boundaries for broadcasting of terrestrial signals in host country language are falling. For example, Danish viewers can get Norwegian programming via satellite or cable, and are generally happy to experience their entertainment in either language.

So now, what was once a geographic space protected by both language and state boundaries, is a smorgasbord of programmatic options in a variety of languages - where the best, most engaging programming will "win" the viewer eyeballs regardless of country of origin or language. This is particularly so in the younger demographics.

Versioning of USA programming is very popular across Europe due to the high production value; but even this is likely to give over to broadcasting such shows in English without the versioning in the coming years. Most important - beyond the language issue - is that good "entertainment engagement" and cultural stories/legends with insightful content can much more easily be adapted or taken into new territories and markets than ever before.

A great example - "*Beowulf*" originally an Anglo-Saxon poem and cultural legend - was brought to the big screen in English and other languages, along with multi-platform components over the Internet. There is BIG potential for indigenous/traditional story-telling in new concepts/platforms from Irish, Spanish, Scandinavian, Romanian or other.

British / French / German Dominance in Programming

The big broadcasters and production houses of Britain, France and Germany dominate the European TV landscape today. And they are also dominating the web-based programming emanating from this region.

This region is also strong in the production and export of Reality TV & Formats for Game-type shows. Of course they have a lower budget than drama, but they are also successfully being exported to USA, Canada, Australia, South America, etc.

Despite the dominance of these 3 countries in the region there is however, still good potential for coproduction with North American or Asian-Australian initiatives, especially in new growth areas of old Eastern Europe: Poland, Czech Republic, Ukraine.

What is unclear is if there is an "event" viewing culture as is prominent in North America for other than sporting events. The Digital Video Recorder (DVR) is also really starting to make headway into these markets.

VoD is growing in popularity and with the huge growth in broadband/IPTV as mentioned above, this region will soon be accessing productions from around the world in a subscription or pay-per-view model, and quite possible through Internet search engines rather than traditional broadcasters. Remember, Joost originates from Europe!

Public Broadcasters Dominate 60+ Age Group

Unlike North America, the biggest broadcasters such as BBC, ZDF, and the leading broadcaster in each country are public broadcasters. They have been defining the evolution of TV in their countries and "protected markets" for decades.

Indeed most of the Public Broadcasters have their biggest audience segment at the Age 60+ group as well as the young children on the other end. GenX32 and even the young people are all watching the private broadcasters which generally dominate market share in the 12 - 59 age segment.

These mighty powerhouses are struggling for relevancy with the majority of the population with disposable income. The role of public broadcasters in European society is not that far behind the challenges being faced by public broadcasters in North America.

Programming style, development approaches and story-engagement plus multi-platform extension to the entertainment experience, is markedly different in Europe between the public and private broadcasters.

4.4 Advertising

There were many media organisations making predictions for trends in 2009. These included paid 'twitterers' appearing on Twitter, qualitative analytics becoming increasingly important for advertisers; localised search and localised services growing, mobile advertising increasing; netbooks impacting web page design; social media expanding and advertisers remaining wary of spending in current economic climate. None of these were especially revelatory except perhaps in the confirmation that the industry remained in a highly conservative state of mind following the economic upheavals of 2008. Somewhat more interesting was the assertion that *application programming interfaces* (APIs) and 'mashups' could be indicative of the next major trend in online media, with users becoming able to create their own front-ends to diversely located content. If this turns out to be the case it represents yet another challenge / opportunity scenario for an already challenged industry. One more traditionally optimistic view was that increasing the quality of creative development in advertising was and would remain a key differentiator.

Advertising spend in the US was down in 2009 by almost exactly 10%. Online advertising however continued to grow, partly by expansion and partly by sucking budgets from other areas; Online is expected to have grown by 8.9% in 09 (down from a growth of 11.3% in 08) but is expected to surge by 10.9%, 12.3%, 15.6% through to 2012 before approaching saturation at 13.5% in 2013; these estimates may not however have taken into account the sudden surge in mobile internet applications in 09.

5 Digital Media Research

This Chapter outlines the work done and results achieved by research partners in the context of the SALERO project. Related research carried out by other external groups are also included for comparison and this will enable partners to make adjustment to future research direction as necessary.

5.1 Digital Media Research by SALERO Partners

This section outlines the areas of digital media worked on by the different SALERO research partners and the tools they developed to evaluate the concepts.

5.1.1 Character Animation

FBM-UPF has developed a series of new technology to assist the creation of 3D animated avatars. The bulk of this work consists of a new pipeline for facial animation; once the modeller/animated has created their avatar's face, they can use the Maskle tool to apply the activation-evaluation emotional model to the character. The entire pipeline saves hours of character animation time, and can be used to procedurally animate several similar characters (for example, in a crowd). FBM-UPF has presented the framework for automatic programme generation at the ACM conference in Advances in Computerised Entertainment Technology. The framework is designed to automate the process of animated programme generation, and is capable of creating templates to facilitate the automatic generation of variations on the basic scripted programme. The framework is capable of automatically generating videos using a variety of codecs and for a variety of platforms.

5.1.2 Semantic Modelling, Annotation and Search

JRS and **LFUI** research in the project resulted in a set of formal ontologies to model 3D content and semantic annotations, the semantic workbench, a means to manage ontologies, and tools for annotation and search, both bundled in the Intelligent Media Annotation and Search (IMAS) system. The annotation system is agnostic to the type of resource being annotated and can be used for any file type occurring in a media production, given that distinct visualization components are provided to ease the annotation. The workbench on which the annotation and search system offers a set of services for annotation and search that are supposed to be directly used in media production tools to foster creation of metadata just in time of the creation of new content.

5.1.3 Audio Transformation

UPF work related to intelligent authoring tools, help sound designers and audio content creators easily transform audio in an intelligent way. The Voice transformation software allows the creation of a set of new voices by means of transforming a voice either in real time or a recorded one. It is also possible to transform a synthesized voice generated using a Text to Speech system like the one developed by URL in the context of SALERO. The *Tempo and Pitch* music transformation software allows the user to create new audio material by changing the tempo/speed or pitch of a song. The preservation of timbre in the audio allows, for instance to shorten slightly an audio clip to match with the duration of a corresponding video sequence (imagine an advertisement on TV where the duration is crucial) or even give a more happy mood to a song transforming it to a higher pitch. The work related to source separation, allows remixing songs to give more prominence to or just remove some instruments within a polyphonic mix as if you had the multi-track recording of any piece. This allows customizing the equalization of a song beyond the classical frequency equalizer.

5.1.4 Speech in Animation

DIT research has produced the *Vowel Builder* and *Vowel Player* online character animation tools providing a cross platform authoring and delivery mechanism for interactive, reusable and multilingual content. The system is designed to work with any reusable avatar, and thus separates the authoring and rendering processes. Interactivity is provided by means of simple scripted narratives, and real-time configuration of an avatar (and indeed its language) are possible. In addition, the system allows multiple characters to be authored and rendered in a single pass, providing means for animation of dialogue speech using one asset.

5.1.5 TTS Synthesis

Funitec-URL tools are mainly based on the reduction of costs on creating new synthetic voice for any Text-To-Speech synthesis whether is Domain-Specific or General Purpose system. These tools are being developed along the Core System of the URL synthesis engine. The main value-added feature of Funitec-URL TTS is the possibility to work with the animated audiovisual productions time line being able to parse tagged information, to synthesize the desired speech with the appropriate constraints and to provide time referenced output regarding parameterization about event and lip synchronization. The appropriate input constraints for the synthesis may be any prosodic requirement as intonation, speed of delivery or intensity (stress) of the speech.

5.2 Similar Media Research Projects by External Groups

5.2.1 Character Animation/Automatic Programme Generation

There is very little academic work that is similar to the Automatic Programme Generation framework developed by UPF, although two similar commercial products do exist. *Redboard* (<http://www.redboard.tv>) is a storyboarding and prototyping technology, designed to enhance existing workflows by integrating the traditional skills of the storyboard artist into a CGI workflow; while *Xtranormal* (<http://www.xtranormal.com>) have developed an online system that allows novice users to “make movies”, using several preset combinations of backgrounds and characters. Neither system allows automatic programme generation.

5.2.2 Semantic Media Annotation

The organisation, classification and retrieval of media objects is an ongoing challenge in games and media production. Semantic technologies have been identified as a viable solution to overcome some of the problems in this area¹⁹.

A wide range of multimedia annotation tools^{20,21} already have over functionality to attach ontological annotations to parts of the multimedia content and some over reasoning services on top of them to semi-automatically create annotations based on existing annotations.

*PhotoStuff*²² allows using any ontology for the annotation of images and is available as a standalone desktop application. A web-based demonstrator for browsing and searching with very limited functionality is also available.

The *K-Space Annotation Tool*²³ provides a framework around the Core Ontology for Multimedia for efficient and rich semantic annotations of multimedia content.

*ImageNotion*²⁴ already provides an integrated environment for the collaborative semantic annotation of images and image parts. User tests showed that the use of standard ontologies and tools is not generally suitable, which led to the development of a method where ontologies consist of image notions that graphically represent a semantic notion through an image.

As experience shows most paradigms applied in semantic annotation tools are not suitable for inexperienced users who are typically used to keyword-based tagging and suffer from information overload when confronted with complex interfaces.

In asset management one can see a trend towards the use of semantic technologies for organization and consequently search and browsing of media assets. One example is the *IPV Teragator*, a metadata aggregation and management tool first presented at the IBC 2009. It allows semantic searches and relational browsing of assets. Another application to mention is the Interactive Media Manager from Microsoft, which however has been discontinued in late 2009. Big players in broadcasting such as the EBU also recently released semantic versions of some standards.

5.2.3 Content Search & Retrieval

In the domain of video retrieval, the *TRECVID*²⁵ effort continues to provide the main focus for research groups around the work. *TRECVID* is sponsored by the US government via the National Institute of Standards and Technology (NIST), and is an offshoot of the earlier Text Retrieval Conference (TREC) effort. *TRECVID* provides research groups with access to large video collections, search topics and associated relevance judgements – all vital resources for video retrieval research. For the domain of image retrieval, the *ImageCLEF*²⁶ effort has become one of the most important TREC-style evaluations.

ImageCLEF is the cross-language image retrieval track of the *Cross Language Evaluation Forum* (CLEF), and is the source of a number of important image collections used in research.

Based on the results of *TRECVID* 2009, the trend has continued towards the use of automatic annotation in the better performing systems. Performance using only visual features continues to be very low when compared to textual systems, in both *ImageCLEF* and *TRECVID*. Of relevance to the retrieval tasks carried out in SALERO, there has been some more recent work aiming to retrieve vector images, although the experimental results are again very mixed, bringing home the difficulty of the image and video retrieval problem, when text is not available.

*PHAROS*²⁷ will move search engines right into the 21st century, providing automatic annotation and full multimedia searching on audiovisual content whatever its nature and structure. The *PHAROS* mission is to advance audiovisual search from a point-solution search engine paradigm to an integrated search platform paradigm. This platform will be built on an innovative, open, and distributed architecture that enables consumers, businesses and organisations to unlock the values found in audiovisual content. The *PHAROS* search platform will create a new infrastructure for managing and enabling access to information sources of all types, supporting advanced audiovisual processing, content handling, and management that will enhance control, creation, and sharing of multimedia for all users in the value chain. The impact for the specific audiovisual industry will be to strengthen and extend product and service offerings, integrating technologies and achieving a competitive advantage by integrating solutions addressing the full content management processing chain.

5.3 Emerging Results from SALERO Research

5.3.1 Character Animation

FBM-UPF research into new animation techniques is now being extended to the whole body. Using motion capture technology to record a male and female actor/actress modifying their walk-cycle according to different emotions, FBM-UPF is developing a system for procedural generation of virtual character walking based on emotional aspects.

5.3.2 Media Annotation

LFUI research resulted in the semantic workbench as an extension of already established ontology management tools and frameworks. In addition a set of ontologies have been created that can be used outside of the project such as ontologies for the modelling of virtual characters or relationships between assets in media production.

5.3.3 Voice Transformation

MTG-UPF research on Voice Transformation has emerged in multiple applications including real-time installations for museums, software plug-ins aimed from amateur to professional users as well as web applications where the technology is implemented as a web-service to transform sound files provided by any web client. Also the technology has been successfully integrated with the URL text-to-speech system to transform the output synthetic voice to many different characters, including male, female, old, child, robot, alien, monster, clown.

5.3.4 Speech Analysis & Annotation

DIT research into the vowel stress tagging framework has led to a patented method of speech analysis for annotation, query, retrieval and animation authoring that can be performed online. The tagging framework underpins research into the acoustic correlates of emotional speech, while also providing analysis data for online speech corpora visualisation tools. The framework also forms the basis of character animation authoring and rendering tools, which are engine independent and scalable to facilitate dialogue animation authoring.

5.3.5 Content-Based Retrieval

UG work on SALERO has resulted in a number of research and engineering outcomes. The main engineering outcomes have been:

- The development of the video/image retrieval backend system

- The development of the *AspectBrowser* search interface, and corresponding indexing interface
- A collaborative search interface was developed
- A personalised news retrieval system was created which integrated a high-level ontology

Carrying out basic and applied research has been an important part of the UG work over the SALERO project. Main research outcomes include:

- Basic research into how users search when faced with complex search tasks
- Investigations of collaboration in search have been carried out
- Work has been carried out to improve the performance of low-level features in search
- Work has, which is continuing, to improve the automatic annotation of image and video data
- Development of a *Latent Dirichlet Allocation* (LDA) based model for multimedia retrieval was created

This research has produced a number of research publications at various international conferences including ACM SIGIR, ACM Multimedia, the *European Conference in Information Retrieval* (ECIR), among others. These research outcomes are in addition to the previously listed tools which have been created.

5.3.6 TTS Synthesis

The quality of synthetic speech is highly affected by the coverage of the linguistic units in the target domain by the speech corpus used. By using the same database as the *Blizzard Challenge 2008*¹⁸, Funitec-URL was able to achieve improved results on TTS English speech. In addition, a new approach based on *case base reasoning* (CBR) helped to improve prosody modelling, resulting in more natural synthesized speech. The CBR scheme also allows a simple treatment of discrete and numerical attributes (without discretization) and numeric array classes (the parameters to be predicted).

In addition, voice quality has been demonstrated to be useful to be used in expressive speech synthesis, so its use along with prosody will be the next step in the development of improved expressive TTS systems.

Finally, the quality of the generated speech as well depends on the accuracy of the unit selection process, which in turn relies on the cost function definition. This function should map the user perceptual preferences when selecting synthesis units, which is still an open research issue. We propose a complete new methodology for the tuning of the cost function weights by *fusing* the human judgments with the cost function, through efficient and reliable interactive weight tuning. To that effect, *active interactive genetic algorithms* (aiGA) are used to guide the subjective weight adjustments. The application of aiGA to this process allows mitigating user fatigue and frustration by improving user consistency.

5.4 Future Directions for SALERO Research Partners

5.4.1 Character Animation/Automatic Programme Generation

In the fields of automatic programme generation, FBM-UPF is looking towards educational applications. Of particular interest is the possibility to automatically create animated avatars that use sign-language to present programmes such as the news or the weather.

5.4.2 Media Annotation

For LFUI, future work will continue the paths followed in the recent months in SALERO into the direction of motivation of users to contribute semantic annotations, modelling of annotations, and large-scale annotation of media on the Web.

5.4.3 Voice Transformation

Future directions for research within MTG-UPF will be related to voice conversion, allowing transforming the timbre of the voice of a source speaker into the voice of a target speaker, maintaining the prosody

and intonation of the source, a lot of difficulties arise, some of them difficult to solve, but the applications of this research are enormous.

5.4.4 *Speech Tagging*

Future work in DIT will focus on the implementation of the tagging framework in other animation rendering engines (notably *3DSMax* and *Maya*). In addition, integration of online animation and speech corpora visualisation technologies will focus on the potential for multi-lingual language learning applications to be developed using the underlying vowel stress tagging framework.

5.4.5 *Image/Video Retrieval*

For UG, future work in video and image retrieval looks to continue on the current path, with the emphasis placed on automatically detecting concepts from video streams or static images. Automatic annotation still requires the use of low-level visual features however, and so it is very unlikely that the use of this information will disappear, but instead the greater application of machine learning techniques will be applied to the problem. At Glasgow University, forthcoming work is likely to concentrate on interfaces and techniques to improve collaboration in search.

5.4.6 *Text-To-Speech Synthesis*

Future directions for Funitec-URL research will be related to expressive speech synthesis and voice transformation taking into account a target expressivity and speaker, by using prosodic and voice quality parameterizations.

6 SWOT Analysis – in the Light of Current Trends

We are clearly seeing wider ways of ‘consuming’ media, whereby a large amount of cross media productions is made with manual intervention. It is reasonably obvious that the trend we will see next is the automatic cross purposing of media productions.

The SWOT analysis presented in the previous version of the document is still valid in general and is repeated here for easy reference.

Strengths	Weaknesses
<ul style="list-style-type: none"> • RTD in intelligent media (in audio, visual and language domains) capable of leading to advances in media production and cross-media repurposing at reduced costs. • Involvement of creative professionals and users as well as technologists. • Integration of technologies in experimental productions (using standards, web services, etc.). • Up-to-date with regards to FBX standardization Publishing research within the field of character Animation 	<ul style="list-style-type: none"> • It is difficult for intelligent technologies to meet both the quality standards required by high-end media professionals <i>and</i> the speed and simplicity required for UGC and mobile content. • Lack of suitable multimedia standards platforms. • Interoperability may be reduced by lack of appropriate standards. • Experimental productions need to integrate different formats in cross-media contexts.
Opportunities	Threats
<ul style="list-style-type: none"> • Use of intelligent content technology tools to enable low-cost production by a wide range of content creators. • Niche content creation by SME producers. • New business based on cross-media development, user-generated content and IPTV services. • User Group activities to get feedback from outside professionals. 	<ul style="list-style-type: none"> • ICT solutions from other providers and research groups in a fast-moving area. • Low-cost content from outside Europe. • Lack of interoperability due to stop-down standardization process.

It should also be noted that work completed by the SALERO partners covers many areas of the digital media industry and it is difficult to summarize all in a single SWOT table.

7 Glossary

Partner Acronyms:

AM	Activa Multimedia, ES
BLITZ	Blitz Games, UK
DFT	Digital Film Technology, DE
DIT	Dublin Institute of Technology, IE
DLLNI	DTS Licensing Limited (Northern Ireland), UK
FBM-UPF	Fundació Universitat Pompeu Fabra, ES
JRS	JOANNEUM RESEARCH Forschungsgesellschaft mbH, AT
LFUI	Leopold-Franzenzs Universtät Innsbruck, AT
MTG-UPF	Music Technology Group, Universitat Pompeu Fabra, ES
PGP	Pepper's Ghost Productions Ltd., UK
TAIK	Taideteollinen Korkeakoulu, FI
UG	University of Glasgow, UK
Funitec - URL	Universitat Ramon Llull, ES

8 References

This section lists all the references used in the main document.

- 1 <http://www.thewrap.com/article/crack-blu-ray-pot-gold-8564&usg=AFQjCNENHcuMam-FR4rAUwYcqYaN4Yjapw>
- 2 <http://www.homemediamagazine.com/general/home-entertainment-spending-gap-continues-narrow-17343&usg=AFQjCNFRKuni1BEBDpA3epRpNPkJY-CF5g>
- 3 <http://www.digitalcinemareport.com>
- 4 SMPTE Motion Imaging Journal September 2009
- 5 http://www.maximumpc.com/article/news/samsung_predicts_hddpriced_ssds
- 6 <http://www.hitachigst.com/portal/site/en/template.MAXIMIZE/company/pressroom/>
- 7 http://www.theregister.co.uk/2009/08/13/imation_lto5_media/
- 8 <http://www-03.ibm.com/systems/clusters/software/gpfs/index.html>
- 9 <http://prosony.com/bbsccms/asssets/files/mkt/digicinema/pressreleases>
- 10 http://www.digitalcinemareport.com/Regal-Sony-Barco-Christie_NEC-DLP-LCOS-4k-2k
- 11 <http://mkpe.com/publications/d-cinema/reports>
- 12 <http://www.minoru3d.com/>
- 13 <http://www.blu-ray.com/news/?id=2104>
- 14 <http://www.youtube.com/watch?v=89cGA4-yPRw>
- 15 <http://www.natoonline.org/pdfs/PDF%20Digital%20Cinema/NATO%20Digital%20Cinema%20System%20Requirements%20-%20release%202%201.pdf>
- 16 <http://www.cedmagazine.com/SCTE-3D-standards-project-022009.aspx>
- 17 http://www.smpte.org/events/smpte_annual_tech/schedule/05wedam3/
- 18 V. Karaiskos, S. King, R.A.J. Clark, C. Mayo, "The Blizzard Challenge 2008"
- 19 Bürger, T.: The need for formalizing media semantics in the games and entertainment industry. *Journal for Universal Computer Science (JUCS)* (June 2008).
- 20 Obrenovic, Z., Bürger, T. Popolizio, P., Troncy, R.: Multimedia semantics: Overview of relevant tools and resources. http://www.w3.org/2005/Incubator/mmssem/wiki/Tools_and_Resources (September 2007).
- 21 Simperl, E., Tempich C., Bürger, T.: Methodologies for the creation of semantic data In: Miguel-Angel Sicilia (ed.) *Handbook of Metadata, Semantics and Ontologies*, World Scientific Publishing Co., 2009.
- 22 Halaschek-Wiener, C., Golbeck, J., Schain, A., Grove, M., Parsia, B., Hendler J.: PhotoStuff - An Image Annotation Tool for the Semantic Web. In: *Poster Proceedings of the 4th ISWC (2005)*.
- 23 Saathoff, C., Schenk S., Scherp A.: KAT: The K-Space Annotation Tool. In: *Proceedings of the SAMT 2008 Demo and Poster Session, (2008)*.
- 24 Walter, A., Nagypal, G.: IMAGENOTION - Collaborative Semantic Annotation of Images and Image Parts and Work Integrated Creation of Ontologies. In: *Proceedings of 1st Conference on Social Semantic Web (CSSW)*. (2007) 161-166.

- 25 <http://www-nlpir.nist.gov/projects/trecvid/>
- 26 <http://imageclef.org/>
- 27 <http://www.pharos-audiovisual-search.eu>
- 28 SALERO deliverable D.10.3.3: Digital Media Technology, Research and Production Trends Report. Version 3
- 29 http://www.hboimagine.com/#/the_affair/ on 16 November 2009
- 30 Mobile Art Lab and the Phone Book:
http://www.mobileart.jp/phonebook_en.html<http://www.everydayux.com/2009/11/09/a-peek-at-the-future-of-interactive-storytelling>
- 31 Finnish Film Foundation: Audience figures for Finnish films in 2009 (on 16 November 2009)
- 32 IBC 2009 Conference and Exhibition, Amsterdam RAI, September 2009: <http://www.ibc.org>
- 33 Wikipedia (http://en.wikipedia.org/wiki/3_D) on 16 November 2009
- 34 UBI (UrBan Interactions) Research Program information and publications are found on <http://www.ubiprogram.fii>
- 35 Finnish Broadcasting Company YLE Areena IP TV service: <http://areena.yle.fi>, YLE Cultural Fitness project: <http://kulttuurikunto.yle.fi/inenglish>
- 36 Massachusetts Institute of Technology; MIT Media Lab and MIT Tangible Media Group:
<http://www.media.mit.edu/>, <http://tangible.media.mit.edu>, Reactable Systems / Pompeu Fabra University: <http://www.reactable.com>
- 37 Korpilahti, T. 2009. From media installations to non-material media art: A study of the different characteristics of physical media installations and non-material media art
- 38 ACM International Conference on Multimedia, Beijing Hotel, Beijing, China October 19-23, 2009:<http://www.acmmm09.org>
- 39 Media facades Facebook group <http://www.facebook.com/pages/Media-Facades-Festival-Europe-2010/152998564085>, Media Facades festival site from 2008: <http://www.mediaarchitecture.org/mediafacades2008/>), Association for new media culture – m-cult <http://m-cult.net>
- 40 From 1 January 2010, TAIK will be known as Aalto University School of Art and Design
- 41 City Wall: <http://citywall.org/>, IPCity project: <http://www.ipcity.eu/>, Ubiquitous Interaction research group at Helsinki Institute for Technology: <http://www.hiit.fi/uix>
- 42 Tikka, P. 2008. Enactive Cinema: Simulatorium Eisensteinense, Helsinki: University of Art and Design Publication Series. ISBN 978-951-558-272-0,
<http://www.oblomovies.com/piatikka/CV.html>
- 43 Best, A. 2009. Media Lab doctoral research proposal: Playful Desires – Physical human-computer interaction through play and performance, September 2009