MANAGING MULTIMEDIA CONTENT
A TECHNOLOGY ROADMAP

Report presenting an emerging technology roadmap for a long-term plan of multimedia indexing and retrieval systems.

08 April 2005

© Wole Oyekoya
Sira Group
o.oyekoya@adastral.ucl.ac.uk

Abstract

The management (storage, retrieval and processing) of multimedia data (audio, image, text and video) is becoming more crucial in this information age. Integration of digital devices such as digital cameras, mobile phones, PDAs and computers has contributed to the critical need for automated multimedia indexing and retrieval of relevant information. This report examines the state of the art in multimedia retrieval and outlines a roadmap with the aim of directing research strategies that can lead to industrial applications. The overall objective of the exercise is to develop a road map and supporting documentation that will provide the basis for developing technology strategy to identify ways of exploiting multimedia technology in order to develop a leading business position. Also highlighted within this study is the continual change in defining multimedia as it revolves around changes in industry, in patterns of communication, or in broader politics and culture and are bound up with various optimistic and pessimistic visions of the Information Society. Many of these take as their starting point the anticipated convergence of different industries, facilitated by innovation and dissemination of converging technologies. The issue of uncertainties in user responses and acceptance of emerging technologies are ignored, yet future visions simply concentrate on technological potential and supplier's deployment processes.

CONTENTS

TABLES

FIGURES

© Wole Oyekoya

© Wole Oyekoya

1 INTRODUCTION

Multimedia is defined as the technological resource to create or present and control communication or information in a variety of media through an integrated channel. This is often described in terms of the convergence of computer, telecommunications and television technology, as it develops to allow digital storage, manipulation, transmission and reproduction of any media type. Digitisation of multimedia has brought about a massive increase in the storage of the data. Many everyday life activities result in the accumulation of huge amounts of data containing different kinds of information (text, pictures, audio, videos, etc.). The goal of Information Retrieval technologies is to allow one to make an effective use of such data. Storage devices seem to be growing fast and coping with this demand but it has also meant data is sometimes stored and almost never retrieved. Mostly, this is because users either forget possessing this data or cannot locate it when it's needed. This is applicable in all areas of multimedia (audio, video, text and images). Figures 1 and 2 below describes the extent of the desire for computer and Internet access in the USA.

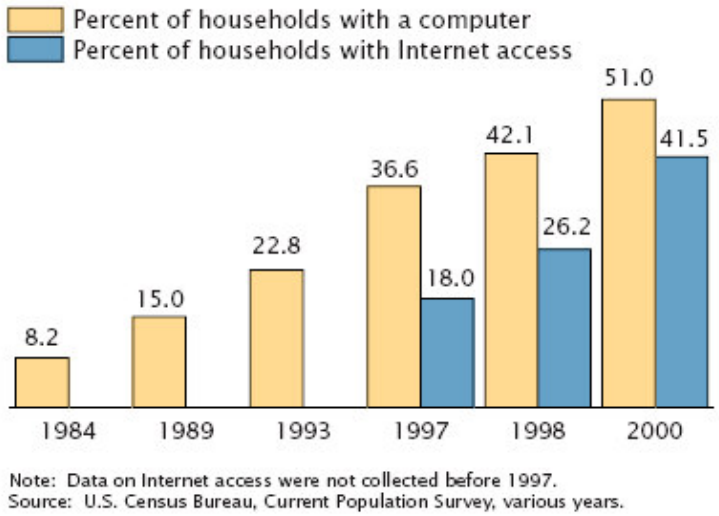
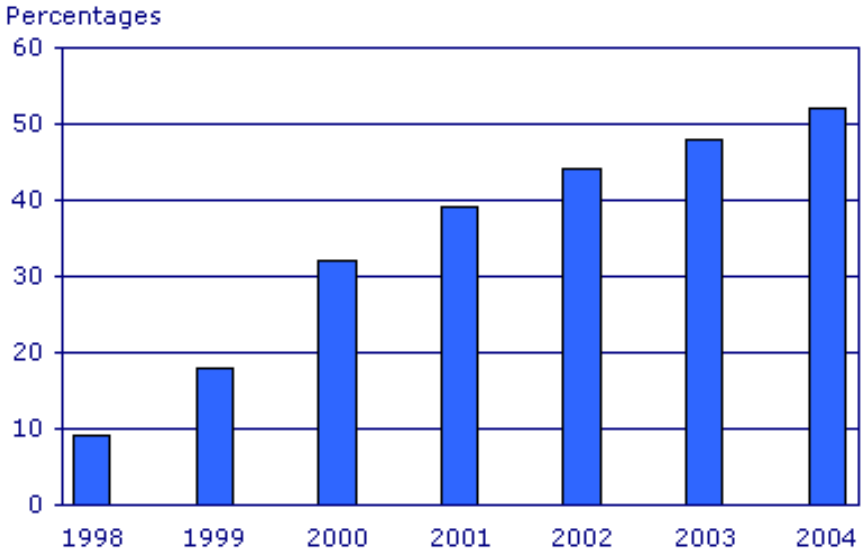


Figure 1: Computers and Internet Access in the Home - 1984 to 2000



Source: Office of National statistics

Figure 2: UK Households with home access to the Internet, July to September (3rd quarter)

© Wole Oyekoya

Text retrieval is now a relatively easy technology due to the advancement of database and storage technology. Text retrieval engines, commonly known as search engines, return a list of documents for a query on a network or local database. Though relevance ranking is still dependent on individual preferences, it is more dependent on the understanding of the language of search. Texts can be extracted from hand-written manuscripts, scanned documents, etc.

Images play an increasingly important part in the lives of many people. The increase in the use of the digital camera has been impressive. The growth in sales has been helped by rapid technological progress and the sales of digital cameras have overtaken standard film cameras. The total sales figure for the six major Japanese digital camera manufacturers, for their own brands of cameras was 40.45 million units for 2003, an increase of 71% from the previous year, and forecasts for 2004 are 61.5 million units, a 52% year-on-year increase. According to figures published in January 2004 by the Camera & Imaging Products Association (CIPA), the global shipment forecast for 2004, including both shipments within Japan and exports, is 60.9 million units, up 40.3% from the previous year. CIPA data excludes production by foreign firms, but is considered the industry standard since Japanese companies command a global market share of about 90 percent. This growth in sales has a number of repercussions in the area of the storage and retrieval of images. The ability to view the image on the camera screen at the time the shutter is released and after it has been captured means that the majority of images recorded are exactly what the user wants. People are taking more photographs and it is envisaged that households will want an automated and efficient method to store and retrieve these images. The integration of digital cameras with mobile devices (such as mobile phones, PDAs and laptops) has added to the critical need for automated photograph management.

Audio retrieval technology poses as difficult a challenge as video but does not have as much network and storage demand. Texts can be extracted (through recognition processes) from speech recordings and music/audio files. The main characteristic of such texts is that they contain errors, hence a certain percentage of words said in a speech recording are transcribed incorrectly. Television programmes currently have to train human experts to re-read information back to a speech recognition software to achieve subtitles during live broadcasts. This is evident in the errors and delay.

Video retrieval is perhaps the most difficult to process, as it encompasses still images (video frames), audio and even textual recognition processes.

Efficient management techniques of audio-visual content can be successfully applied in a number of fields, which includes but not limited to UK government agencies' data collection, security, medical imaging, entertainment (media and content), and defence. However, the home user will be a major player in determining the market for this product as shown by the sort of tasks that is being done on the Internet (Figures 3 and 4). The user wants to be able to use this medium for retrieving not just information but relevant and personalised information.

© Wole Oyekoya

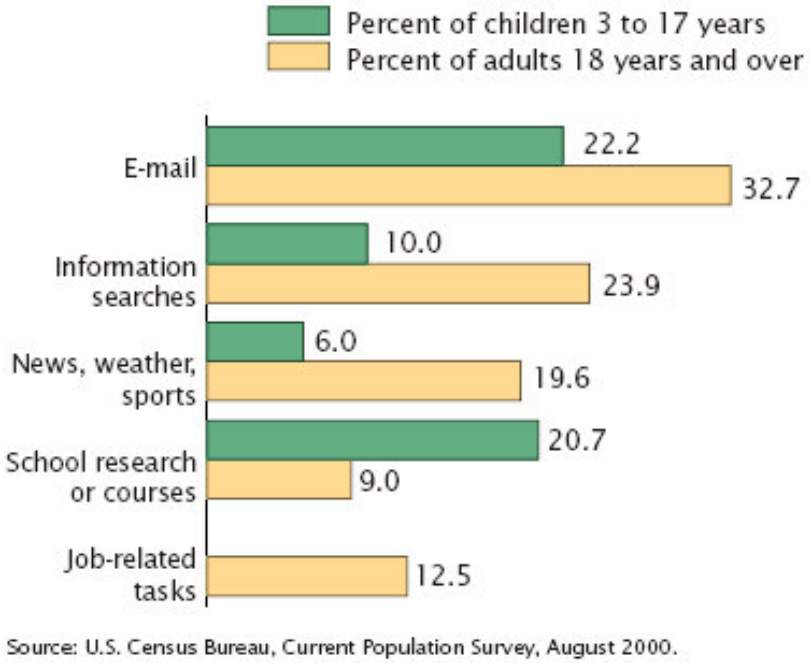
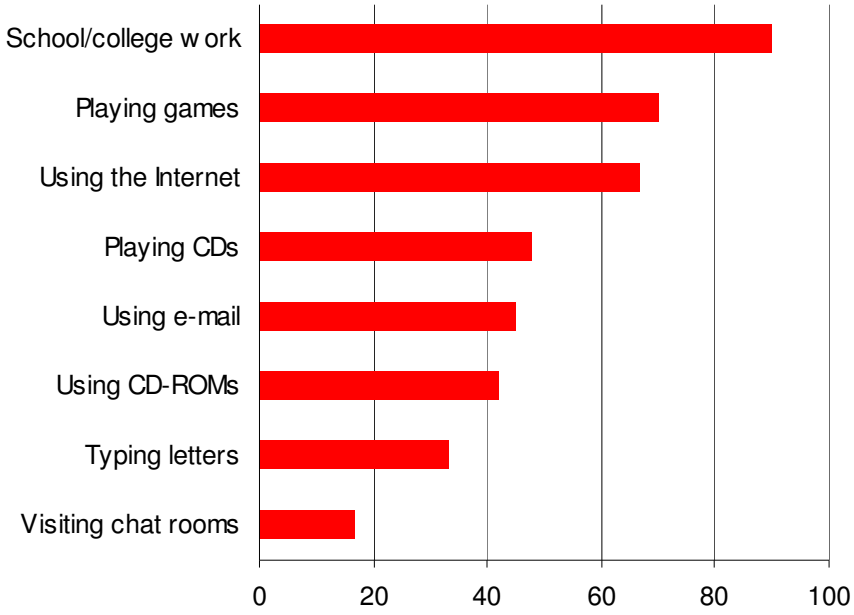


Figure 3: Adults and Shildren Using the Internet for Specific Task - August 2000



Source: Young people and ICT, Department for Education and Skills

Figure 4: Activities undertaken on the computer at home by 11-18 year-olds1 (Percentages) Autumn 2002, England

[1] More than one reason could be given.

© Wole Oyekoya

2 MARKET

2.1 Trends and Drivers

The wide variety of current and past projects reflects the general situation. Trends of all aspects of our life influence the progress of multimedia data management systems: political, environmental, social, technological and economic.

2.1.1 Technological

Probably the most obvious trend is the vast increase in the sales of digitised multifunctional devices. The mass take-up of consumer digital photography creates new demands for handling and archiving the generated materials. Integration of these new technologies creates new challenges in industrial applications. The ubiquity of digital multimedia devices opens up a new era.

2.1.2 Economic

The economic impact of new trends should not be underestimated. Fast and reliable access to information will mean that organisations and individuals can produce at maximum efficiency, which in turn has a positive effect on the economy. New multimedia systems, especially where they include the development of new infrastructure, involve huge investment. This initial investment can be an enormous influence on future development, by institutionalising particular standards, protocols and platforms on the one hand, and favouring particular knowledge and skills, suppliers and diffusion channels.

2.1.3 Environmental

Relatively, the positive environmental impact of multimedia systems will be immense. Automatic systems that are capable of handling digital data will improve the overall performance of humans and maximise the usage of relevant data in decision-making processes. Physical storage could become obsolete due to adoption of e-business technologies. Many educational institutions, businesses, organisations (large/small), financial institutions and individuals have supported the development of multimedia systems to enhance learning, accessibility, e-commerce and to also support such framework.

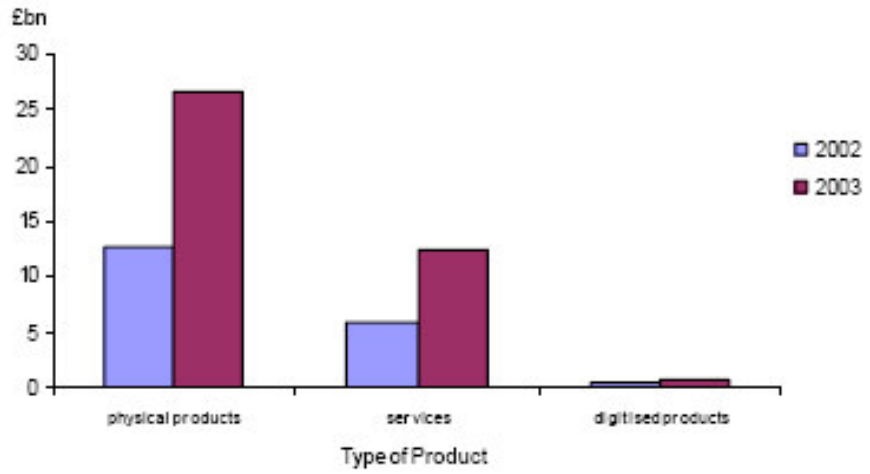
2.1.4 Social

There has been an increase in demand for entertainment, as lifestyles and working patterns change over the years. The explosion of digital content across several platforms such as interactive television, the Internet and mobile communications play a role in the change in societal attitudes. The needs of an increasingly aged and affluent society will be a major driver for improved healthcare, with a greater requirement to image patients for extended periods during treatment and to screen for predisposition to disease, necessitating cheaper, less-invasive and more compact imaging methods. Stewart and Williams discusses the co-evolution of society and technology [7]. Social and economic changes have profound effects on technical innovations. The emergence of life long learning, changes in the international distribution of industry, an evolving military agenda dealing with new terror threats, and an increasing multilateral global cultural exchange all play a part in shaping the future. Users are slow to catch up with technology possibly because of their lack of knowledge of its capabilities are frequently unimpressed by descriptions of expensive new technology. Thus, social learning becomes an important part of the innovation process, especially when it is fed back to suppliers through the market or direct contact.

2.1.5 Political/Organisational

The greater emphasis in society on security and crime prevention will also drive political agendas on the need for compact, relatively cheap, non-intrusive imaging devices able to detect weapons, explosives, drugs and identify people. The figure below shows the take-up of e-business by UK organisations.

© Wole Oyekoya



Source: Office of National Statistics

Figure 5: Value of sales over the Internet by UK non-financial sector businesses

These five themes are not independent and have interdependencies between them. For example, it takes a huge and slow effort to develop markets for new technology. The process of making and demonstrating the possibilities of the innovation, developing new standards and encouraging complementary products and services is often only available to the largest organisations. Unless they can attract the attention of larger partners, smaller organisations are often left to fit new technology into older standards, and waiting for changes in complementary technologies (technological), infrastructure, regulations (political/organisational), skills (economic) and priorities (social/political). These problems are particularly acute in relation to multimedia technology (which involves integration of digital devices and services such as the mobile phone, digital camera, multimedia content delivery and network access). Each respective technology requires significant investment in the infrastructure by the manufacturer or service provider. Justification of the investment is uncertain, particularly since the value of participation often depends on the number of other players linked to the service. The most rational decision by the consumer may be to wait and see - particularly where there are competing standards that may lead to collapse of the project overall. Indeed one of the key political and organisational factors shaping these developments is the activity of standard setting bodies, and the attempts by suppliers to create de facto industrial standards by aligning the expectations and behaviour of other players.

2.2 Copyright

Copyright is an unregistered right and comes into being as soon as a copyright work is created. Typically, copyright prevents a protected work from being reproduced and/or distributed without the copyright owner's consent. To be a copyright work, the work must be original and except in the case of artistic works, recorded. The range of things covered by copyright is very wide. Written works of all kinds and in any media, computer programs, drawings, plans for a building, tables and forms, letters, documents, correspondence, songs and music, recordings, television programmes, broadcasts, films and photographs all attract copyright protection. Copyright gets international protection through a series of complex international treaties. However, not all material that is protected in the originating country as copyright is protected at all or to the same degree in other countries.

As information access becomes quicker and more relevant, copyright protection will be the biggest issue that will be debated in the years to come and perhaps the biggest threat to expansion of multimedia technique. One of the major reasons behind copyright remaining an unregistered right is due to the impossibility of its management, however even this technology may change due to improvements in search mechanisms and watermarking techniques.

© Wole Oyekoya

3 PRODUCT/SERVICES

Market forces determine the products that become commercial. The wide variety of current and past projects reflects the general situation. The current big application trends in Multimedia include the ever-increasing popularity of the Internet, which embraces nearly all multimedia technologies/application areas and rapid advances in enabling technologies to support ever-increasing need for Multimedia.

The review of current and past research projects (incl. network of excellence programs), outlined in the appendix gives an idea of industrial perspective on the direction the technology is heading. It also identifies the organisations that are involved in the research projects and network programs.

3.1 Current and Past Projects

A number of WWW search engines (images.google.com, multimedia.lycos.com and gallery.yahoo.com) can seek images/audio/video from the Internet but they are keyword based and the validity of results are dependent on individual annotation of their respective captions and textual descriptions.

Presented in the table below is a list of retrieval systems identified in [2] [3]:

Table 1: Current retrieval systems

Retrieval System	Organisation
QBIC	IBM
Virage	Virage Inc.
RetrievalWare	Excalibur Technologies Corp.
Photobook	MIT Media Lab
VisualSEEk and WebSEEk	Columbia University
Netra	UCSB Alexandria Digital Library (ADL) project
MARS	University of Illinois
Blob-world	UC-Berkeley
CAETIIML	Princeton University
AltaVista Photofinder	AltaVista (developed at DEC Research Lab)
AMORE	C & C Research Laboratories NEC USA, Inc.
Berkeley Digital Library Project	University of California Berkeley
CANDID	Los Alamos National Lab, USA
C-bird	Simon Fraser University, Burnaby, B.C., Canada
CBVQ	Columbia University, NY
Chabot	University of California, Berkeley, CA, USA.
CHROMA	University of Sunderland, UK
Compass	Centre for Scientific and Technological Research, Trento, Italy
Diogenes	University of Illinois at Chicago
DrawSearch	University of Bari, Italy
FIDS	University of Washington, Seattle, WA, USA
FIR	Esprit IV project FORMULA
FOCUS	University of Massachusetts, Amherst, MA
FRIP	Yonsei University, Korea
ImageFinder	Attrasoft Inc.
ImageMiner	Univerity of Bremen, Germany
ImageRETRO	University of Amsterdam, The Netherlands
ImageRover	Boston University, MA
ImageScape	Leiden University, The Netherlands
iPURE	IBM India Research Lab, New Delhi, India
Jacob	University of Palermo, Italy
KIWI	INSA Lyon, France

© Wole Oyekoya

LCPD	Leiden University, The Netherlands.
MetaSEEk	Columbia University, NY, USA
MIDSS	Purdue University, Indiana
MIR	University at Buffalo, NY, USA
Picasso	University of Florence, Italy
PicHunter	NEC Research Institute, Princeton, NJ, USA
PicSOM	Helsinki University of Technology, Finland
PicToSeek	University of Amsterdam, The Netherlands
Quicklook	CNR Institute of Multimedia Information Technologies, Milan, Italy
RETIN	ENSEA/University of Cergy-Pontoise, France
Shoebox	AT&T Laboratories, Cambridge, UK (developed at Olivetti and Oracle Research Laboratory)
SIMBA	Freiburg University, Germany
SMURF	Utrecht University, The Netherlands
SQUID	University of Surrey, UK
Surfimage	INRIA, Rocquencourt, France
SYNAPSE	University of Massachusetts, Amherst, MA
TODAI	EPFL, Switzerland, Halmstad University, Sweden
Viper	University of Genova, Switzerland
VP Image Retrieval System	University of Tokyo, Japan
WebSeer	University of Chicago, Illinois, USA
WISE	Department of Computer Science, Stanford Univerity

Rui, Huang and Chang's review [2] also identifies and suggests future research directions, two of which are including human in the loop (personalisation) and the need for combining CBIR systems with the World Wide Web (integration), which are the main thrust of this report.

4 TECHNOLOGY

Technology provides the principle means by which the required improvement to the multimedia information retrieval will be achieved.

4.1 Technological Directions

It's only in the last few years that smartphones - and their underlying cellular networks - have evolved to the point where audio and video entertainment can be downloaded to the user's handset. The technology behind such mobile content has allowed all four of the main GSM networks in the UK to have ring tones and music files, along with movie, TV and classic comedy clips, available for download on their mobile Internet portals. Audio and video downloads really come into their own, however, when downloaded across 3G networks, with its capability to approach 512K broadband speeds, which is why 3G carriers around the world are keen to offer such services and so boost their revenues. As the flow of digital visual data from these devices increases and is transmitted over the network for storage, retrieval mechanisms must be capable of handling the amount of data efficiently. Existing systems are capable of retrieving archiving material according to date, time, location, format, file size, etc. However, these systems cannot attach semantic attributes without the need for manual intervention. The ability to retrieve images with semantically similar content from a database is of utmost importance. The growth in the number and complexity of image and video collections has meant more efficient methods of storage and retrieval are needed. This has prompted significant growth in research into techniques of automatic indexing and retrieval.

One of the major issues in information searching is the problem associated with initiating a query. Indeed lack of high-quality interfaces for query formulation has been a longstanding barrier to effective retrieval. Users find it hard to generate a good query because of initial vague information (i.e. I don't know what I am looking for but I'll know when I find it). Eye tracking and other natural

© Wole Oyekoya

methods present an adaptive approach that captures the user's current needs and tailors the retrieval accordingly. The provision of relevant and personalised/customised content is key to maximising this revenue generating capability and systems that can anticipate human behaviour may serve as a new source of information that can guide search and retrieval.

Every company's dream is to make a technology pervasive in our everyday life. The realisation of making users dependent on a product without being conscious of the product/service is important in any product's lifecycle. Basically, integration of digital devices and convergence of technologies provides the key to successful pervasive technology. Perceptive interfaces (e.g. eye tracking) will provide a personalised medium for achieving this goal.

4.2 Technology Foresight

Our move towards a personal digital environment will come from our ability to carry a growing number of simple and capable devices, with fully connected technology at the scale of credit cards, glasses, spectacles, wristwatches and so on. The thrust is towards providing intelligence in anything that is most convenient in our personal environment. One likely development of the personal digital environment is the mobile phone. It starts with a taking a picture of a scenery and obtaining information back on the scene but quickly moves on to devices that can actually forewarn users based on location information even before the user requests the information.

The need for analysing the emergence of the convergence of technologies (incl. information and cognitive technologies), and the assessment of the implications of the converging technologies (CT) was highlighted in a report by an expert group to the European Union, which made 16 recommendations [5]. Perhaps, the most crucial of these recommendations is that the Commission should integrate a CT dimension in sixth framework programme (FP6) calls (in particular in the thematic priorities of nanotechnology, life sciences, information technologies, social sciences and humanities). The flagship for the European Union's involvement in research is the FP6. Drawn up by the European Commission, it was adopted by the European Parliament and EU governments in June 2002. It runs from 2002 to 2006 and has a budget of €17.5 billion over the full period, a 17% increase on its predecessor (FP5). The European Union presently uses multidisciplinary networks of excellence and integrated projects to facilitate integration of technologies [6]. A Network of Excellence strengthens the scientific and technological excellence on a particular research topic. Integrated Projects are objective-driven and conceived either to improve competitiveness or to tackle major societal needs. See appendix for a list of typical projects and networks of excellence.

4.3 Research Challenges

1. Multimedia Indexing and Retrieval using Features: Feature extraction is the basis of efficient retrieval. Features may include text-based features (keywords, annotations, etc), visual features (colour, texture, shape, etc) and audio features (estimation of the beats per minute in a music file, etc). Improvements in processing techniques are needed to achieve desired results.

2. Perceptive User Interfaces: Presently, current research focuses on the importance of building automatic systems that excludes the user. The use of a fully automated system without human interaction is in many applications neither possible nor sensible at this time. Although we humans have poor attention spans, we do have a remarkable ability to recognise objects based on very limited information. Decisions often have to be made in a complex scenario according to a certain context in a certain situation. It is therefore often necessary to include a human expert in the loop of data acquisition and decision making. New interactive user interfaces capable of capturing this knowledge have to be found and implemented. Multimodal interfaces such as eye tracking (visual input) and speech recognition are two important methods of including an expert/knowledgeable user within this process. Human behaviour depends on highly developed abilities to perceive and interpret visual information and provides a medium for the next generation of multimedia retrieval interfaces. If the computer can correctly interpret the user's natural behaviour, it will be

© Wole Oyekoya

able to anticipate the user's objectives and retrieve audio, images and video extremely rapidly and with a minimum of thought and manual involvement.

3. Hardware: The conjunction and digitisation of different media like digital cameras on mobile phones, digital television, DVD, digital radio, the Internet palms and laptops with growing computing and storage power, leads to a diverse information, which requires efficient search strategies. The convergence of telecommunications, information processing and content technologies has driven the need for integrating these digital devices.

4. Security: Perhaps, one of the biggest issues facing information access is network security (inc. local, Internet and mobile networks). Increases in network speeds and bandwidths will mean requirement for secure access. However, multimedia researchers have to tackle the security as well, as it might strengthen any advancement in network security. Digital Watermarking is a widely researched solution with applications on images and video. The human capability to make judgements form limited information can also be used to improve security

© Wole Oyekoya

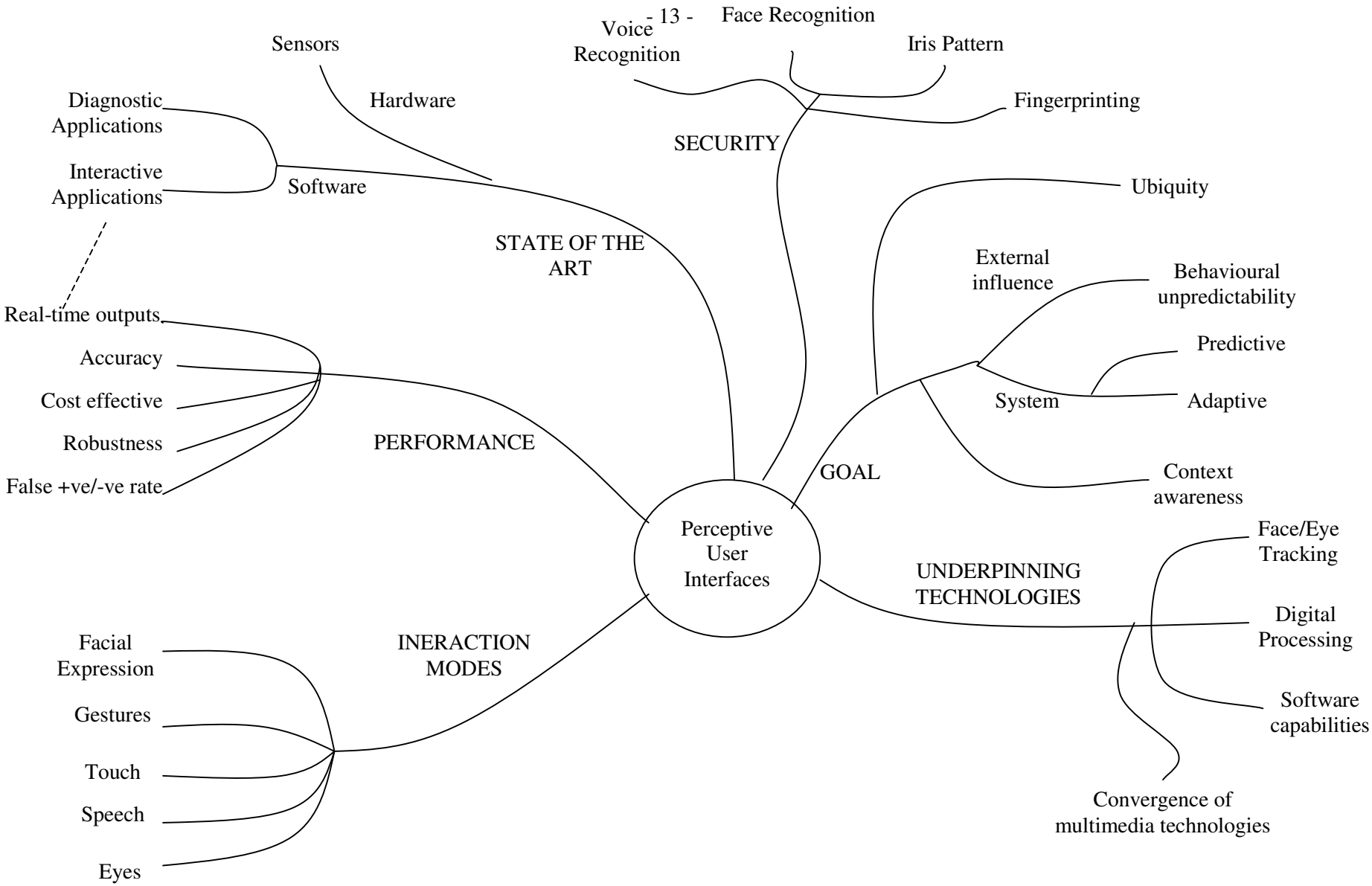


Figure 6: Perceptive User Interfaces

© Wole Oyekoya

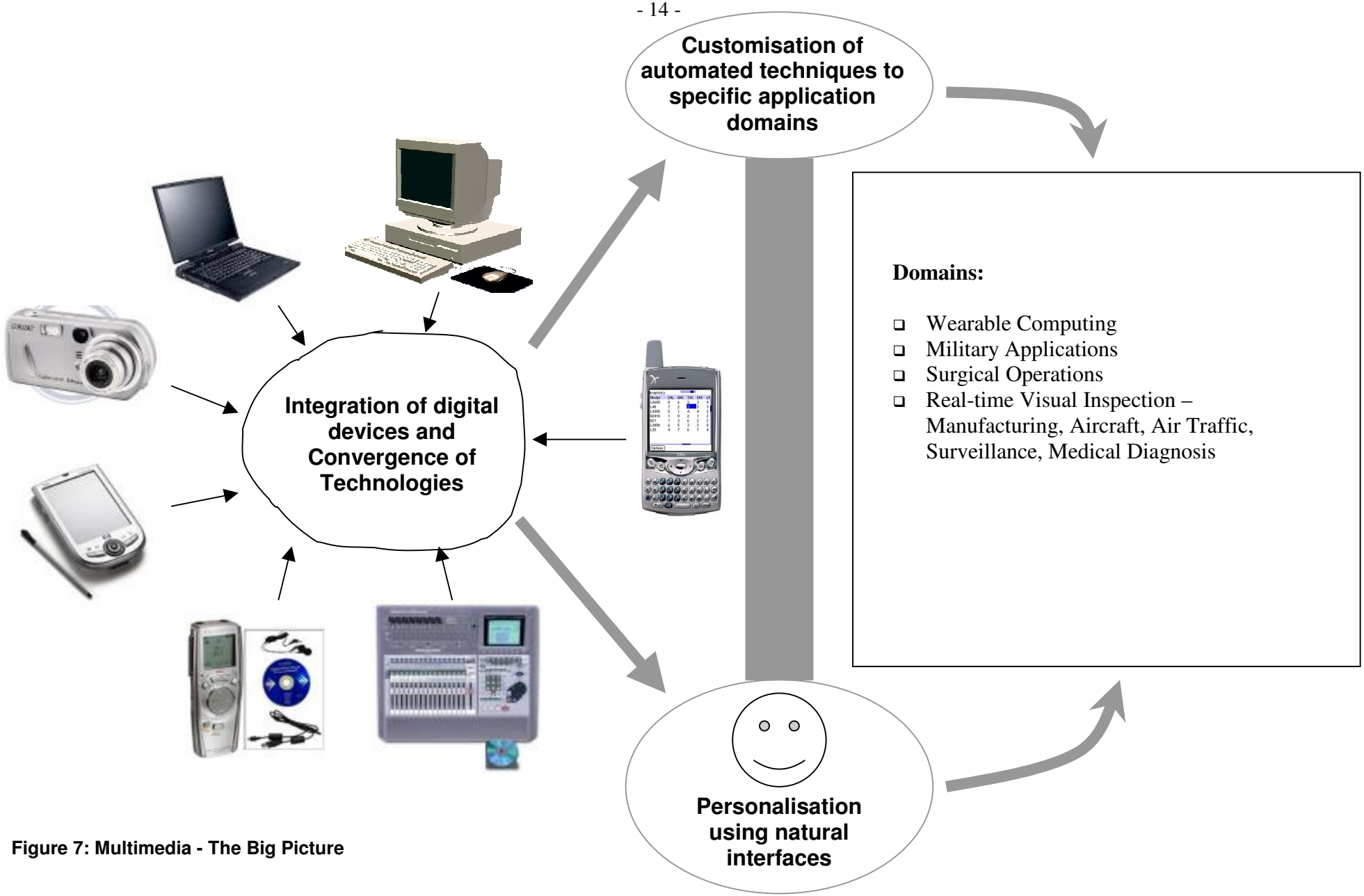


Figure 7: Multimedia - The Big Picture

© Wole Oyekoya

5 RESOURCES

5.1 Funding Strategies

Research Councils such as EPSRC are a source of funding for research in universities. They jointly fund collaborative projects with industrialists in order to develop new technologies to drive innovation in respective key sectors. EPSRC invests in the future of industry and supports post-graduate training in the physical sciences, engineering and technology disciplines to ensure the ongoing supply of experts from academia into industry. They can put organisations and institutions into contact with the university with the relevant expertise to develop the technology relevant to industry.

The Multimedia Demonstrator Programme (MMDP) [1] was the first new programme to be brought forward under the Information Society Initiative (ISI) with DTI support of £7 million and matching funding from industry. Multimedia technology was at its infancy at this stage. It was launched alongside the Information Society Initiative in February 1996. Multimedia covers the full range of information and communication technologies, including video conferencing, CD-ROMs, the Internet, mobile communications and interactive kiosks. The MMDP intended to encourage UK SMEs (Small and Medium Sized Enterprises) to gain greater awareness of the benefits of multimedia technologies, to exploit them in order to enhance their business performance and to stimulate the growth of a healthy multimedia supply sector. There were two competitive calls for proposals from which a total of 51 projects have been supported involving some 199 organisations. A later targeted ad hoc call resulted in support for a further 2 projects involving six partners. Almost £4.5 million of the budget had been spent by March 2000 with £1.7 million forecast to be spent between the financial years 2000/ 1 and 2002/ 3. The final DTI expenditure on the programme is estimated to be £5.5 million.

According to EPSRC, the Media and Content sector, even without the publishing industry, has a turnover of over £18 billion from around 8,000 companies, many of which are small-to-medium size enterprises. It was estimated, in 2002, that 150,000 were employed in the sector plus an additional population of 50,000 freelancers. Computer games are an important component of the industry. Worldwide, it is worth £20-25 billion per annum. The UK is particularly strong in this area, representing around 50% of the total market for computer games development. Characteristic of this industry is the short-term nature of many strategies, particularly in bringing new products to market. Where speed is so critical in securing competitive advantage, the research challenge is to deliver innovation at a commensurate pace. EPSRC currently funds 198 research projects at a cost of £47 million in the media and content industry. Examples of industrial companies collaborating on research projects in Content and Media are: British Broadcasting Corporation; Geneticxchange Inc; GlaxoSmithKline; Hewlett-Packard Company Inc; IBM United Kingdom Ltd; Nucleus Digital Broadcasting Ltd; SoftSound Ltd; Switch Digital (London) Ltd; TTPCom Ltd; Unicom Communications. The interesting observation amongst these companies is that there are no two companies that are from the same sector. The impact of multimedia cuts across different sectors and enables investments in multimedia research by several companies.

© Wole Oyekoya

5.2 National and European Funding Structure

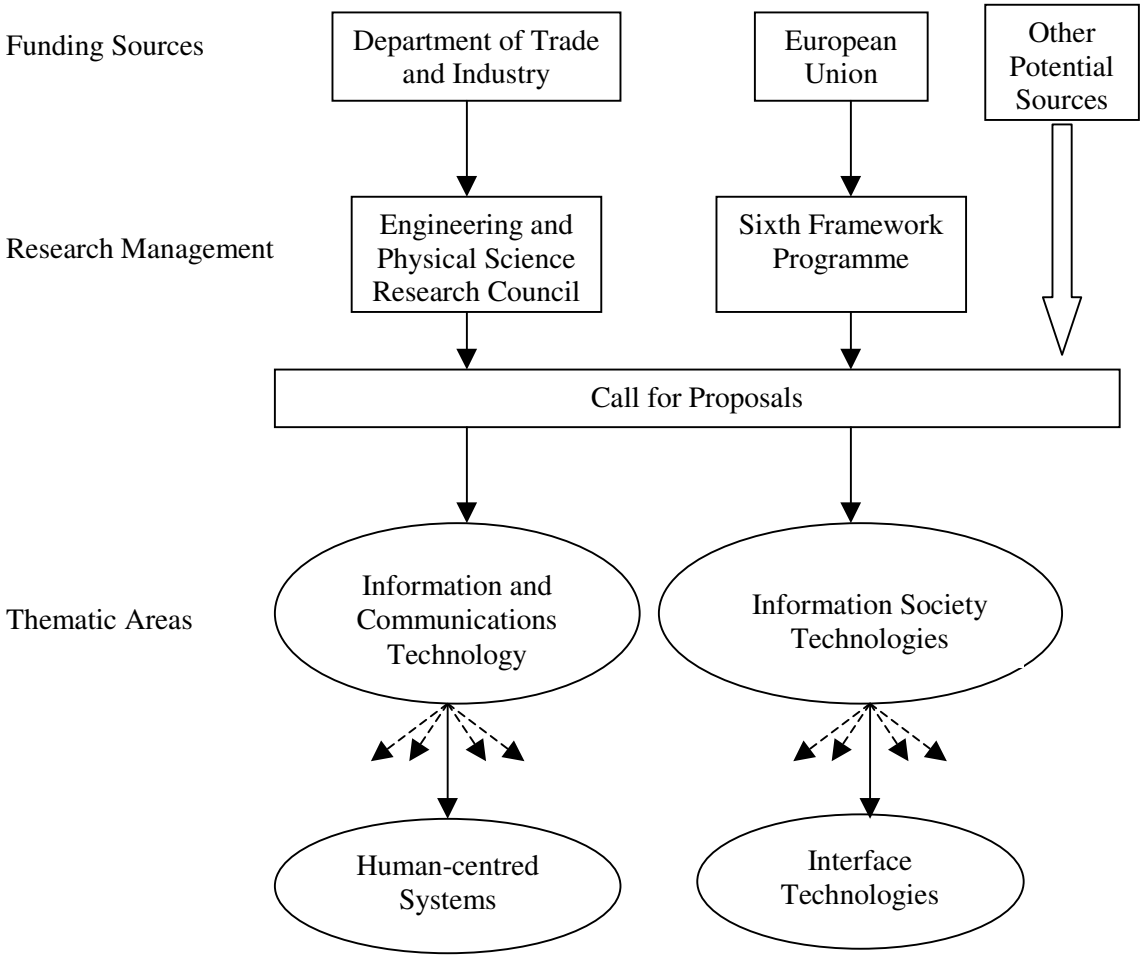


Figure 8: Typical funding structure

5.2.1 Fifth (EC) Framework Programme

Budget Breakdown

Theme 2: Creating a user-friendly Information Society (IST Programme)

Table 2: Key actions for 5th framework programme

Key actions	
Systems and services for the citizen	646 M€
New methods of work and electronic commerce	547 M€
Multimedia content and tools	564 M€
Essential technologies and infrastructures	1363 M€
Research and technological development activities of a generic nature:	
Future and emerging technologies	319 M€
Support for research infrastructures:	
Research Networking	161 M€
Total	3600 M€

5.2.2 Sixth (EC) Framework Programme

The table below presents the calls, the Strategic Objectives that are open in each call, the type of instruments that can be used and the pre-allocated budget per objective.

© Wole Oyekoya

Table 3: Objectives and budgets of key research areas

Areas	Strategic Objectives	Indicative pre-allocated budget €M
Technology Components	Pushing the limits of CMOS, preparing for post-CMOS	75
	Micro and nano systems	85
	Advanced displays	25
	Optical, opto-electronic, photonic functional components	45
	Open development platforms for software and services	55
Integrated Systems	Broadband for all	60
	Mobile and wireless systems beyond 3G	90
	Towards a global dependability and security framework	55
	Multimodal Interfaces	65
	Semantic-based knowledge systems	55
	Networked audio-visual systems and home platforms	60
	Networked businesses and governments	75
	Embedded systems	50
	Cognitive systems	25
	Applications and Services for the Mobile User and worker	60
	Cross-media content for leisure and entertainment	55
	GRID-based Systems for solving complex problems	45
Sectorial Applications	E Safety of road and air transport	65
	eHealth	70
	Technology-enhanced learning and access to cultural heritage	65
	Improving Risk management	30
	eInclusion	30

5.2.3 Multimodal Interfaces Projects Funded under 6th FWP (Sixth Framework Programme)

Table 4: Projects funded under the 6th framework programme

Projects	Description
1. AMI	Augmented Multi-party Interaction
2. CHIL	Computers In the Human Interaction Loop (CHIL)
3. DIVINES	Diagnostic and Intrinsic Variabilities in Natural Speech
4. ENACTIVE	Enactive interfaces
5. HIWIRE	Human Input that Works In Real Environments
6. HUMAINE	Human-Machine Interaction Network on Emotion
7. MWEB	Multimodal Web Interaction
8. PASCAL	Pattern Analysis, Statistical Modelling and Computational Learning
9. SIMILAR	The European research taskforce creating human-machine interfaces SIMILAR to human-human communication
10. TAI-CHI	Tangible Acoustic Interfaces for Computer-Human Interaction
11. TALK	Talk and Look, Tools for Ambient Linguistic Knowledge
12. TC-STAR	Technology and Corpora for Speech to Speech Translation
13. T'N D	Touch and Design

© Wole Oyekoya

Table 5: Key players and funding for multimodal interface projects

Projects	Funding (Cost) (Million €)	Major Companies Involved
1. AMI	8.8 (16.82)	Philips Electronics Netherlands
2. CHIL	15 (23.44)	Diamler Chrysler Germany IBM Czech Trentino Cultural Institute, Italy
3. DIVINES	2.2 (3.63)	Multitel Belgium France Telecom
4. ENACTIVE (NoE)	5	Sony France
5. HIWIRE	1.8 (2.8)	Thales Avionic France
6. HUMAINE (NoE)	4.95	France Telecom
7. MWEB	0.955 (1.24)	W3C
8. PASCAL (NoE)	5.44	Xerox France
9. SIMILAR (NoE)	6.05	France Telecom
10. TAI-CHI	2.35 (3.31)	-
11. TALK	4.40 (5.71)	BMW Germany
12. TC-STAR	10.99 (18.44)	Trentino Cultural Institute, Italy Siemens Germany IBM Germany Siemens France Nokia Finland Sony Germany
13. T'N D	2.22 (3.4)	-

NoE: Networks of Excellence

5.3 People

The trend of resources in the United Kingdom is such that there are a high number of foreign students enrolled in research programmes. This will have a positive impact on the transfer of knowledge into the United Kingdom.

Table 6: Foreign students enrolled in PhD programmes, 1999

Percentage of all students enrolled

1999	No of foreign students
Australia	22.3
Austria	14.6
Belgium	34.1
Canada	18.1
Czech Republic	5.5
Denmark	18.2
Finland	5.6
Italy	2.8
Korea	1.2
Mexico	1.0
New Zealand	8.0
Norway	15.9
Spain	11.7
Sweden	13.9
Switzerland	35.9
Turkey	1.9
United Kingdom	28.8
United States	25.6

Source: OECD (2001c)

© Wole Oyekoya

6 BASIC ROADMAP STRUCTURE

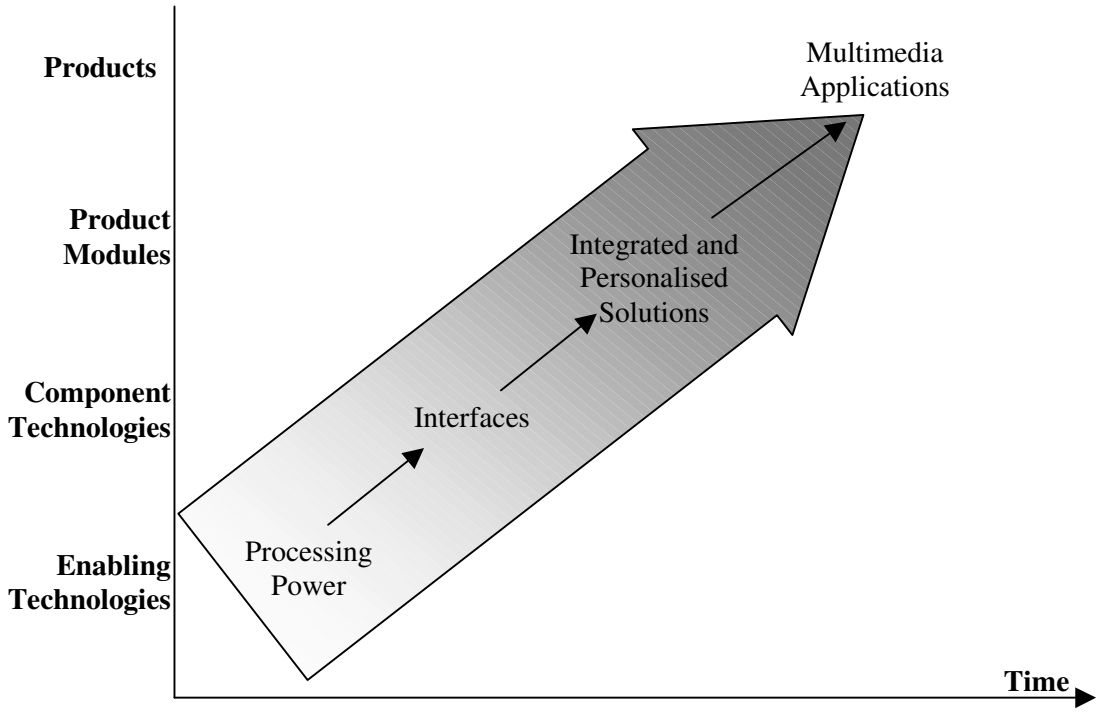


Figure 9: Roadmap Structure

7 CONCLUSIONS

The field of multimedia processing is a technology that will have a strong impact on the economy and our way of life. Fortunately, this opportunity can be utilised by any organisation (large or small). Unlike most research areas, a relatively small start-up company with the right technical expertise and limited funds can get involved in multimedia research, as a result of the continual increase in digital processing power and relative drop in prices. Hence, any company can put themselves at the forefront of this technology. However, it is worth remembering that the provision of relevant and personalised/customised content is key to maximising this revenue generating capability. The driver for any organisation's plan should be integration and personalisation (customisation) of this technology. Perhaps, it might be that the multimedia processing technology is ripe enough to push to the mass market, though the same cannot be said for multimodal interfaces that will be effective in personalising results to respective queries.

© Wole Oyekoya

8 REFERENCES

[1] Department of Trade and Industry Assessment Unit, "*ISI Evaluation"* – Volume 4, Executive Summary 1.

[2] Y. Rui, T.S. Huang, and S.-F. Chang, "*Image Retrieval: Current Techniques, Promising Directions, and Open Issues*," Journal of Visual Communication and Image Representation, Volume 10 (1999), pp. 39-62.

[3] Veltkamp R.C. and Tanase M. (2000) "Content-Based Image Retrieval Systems: A Survey" Technical Report UU-CS-2000-34.

[4] "*Technology Foresight Panels, Expert Papers*", Technology Foresight Summit Budapest, 27-29 March 2003

[5] Alfred Nordmann Rapporteur, "*Converging Technologies – Shaping the Future of European Societies",* Report, Converging Technologies for a Diverse Europe, 14-15 September 2004, Brussels, Belgium.

[6] Working document on foresight in the Sixth Framework Programme*, "Roadmap Foresight in FP6*", version: 27 May 2003.

[7] James Stewart, Robin Williams. "*The co-evolution of society and multimedia technology: issues in predicting the future innovation and use of a ubiquitous technology*", Originally published in Social Science Computer Review 1998, 16(3). Updated August 1999.

[8] European Commission, "*Information Society Technologies: 2003 – 2004 Work Programme*".

© Wole Oyekoya

9 APPENDIX

9.1 Typical Projects

1. COGAIN: Communication by Gaze Interaction

COGAIN (Communication by Gaze Interaction) integrates cutting-edge expertise on interface technologies for the benefit of users with disabilities.

Project details

Project Reference: 511598 Contract Type: Network of Excellence
Project Cost: 2.72 million euro Project Funding: 2.90 million euro

Participant Organisations

University of Tampere, Finalnd**;** De Montfort University, United Kingdom; Risoe National Laboratory, Denmark; The Chancellor, Masters And Scholars of The University of Cambridge, United Kingdom; University of Derby, United Kingdom; Siauliu Universitetas, Lithuania; IT University of Copenhagen Denmark; Danish Centre for Technical Aids for Rehabilitation And Education Denmark; Tobii Technology Ab Sweden; Universitat Zu Lubeck Germany; Technische Universitaet Dresden Germany; Ace Centre Advisory Trust Ltd United Kingdom; Tokyo Institute of Technology Japan; Permobil Ab Sweden; Bispebjerg Hospital Denmark; Universitat Zurich Switzerland; Hewlett Packard Italiana Srl Italy; Universitaet Koblenz-Landau Germany; Danmarks Tekniske Universitet Denmark; Politecnico Di Torino Italy.

2. DRIVESAFE

Development of 3-D Eye tracking device for safer driving and more efficient Web Page production

Project details

Project Reference: 508280 Contract Type: SMEs-Co-operative research contracts
Project Funding: 973908.00 euro

Participant Organisations

Institut National De Recherche En Informatique Et En Automatique, France; Kayser-Threde Gmbh, Germany; Mediascore Gesellschaft Fuledien-Und Kommunikationsforschung Mbh, Germany; Incordia Ab, Sweden; Eurisco International, France; Plan_B Media Gmbh, Germany; Siemens Vdo Automotive Sas, France; Airbus Sas, France; Area 17 Ab, Sweden; Multimedia Ltd, Romania; One Too S.A. France.

3. Improving Display and Rendering Technology for Virtual Environments

The aim of this project is to improve lightweight near-to-the-eye displays and tiled stereoscopic large size displays.

Project details

Project Reference: 4785 Contract Type: Specific Targeted Research Project
Project Cost: 2.87 million euro Project Funding: 1.88 million euro

Participant Organisations

Technische Universitaet Darmstadt, Germany; Stt Ingenieria Y Sistemas, S.L., Spain; Microemissive Displays Limited, United Kingdom; Page And Park Architects, United Kingdom; Elasis S.C.P.A., Italy; Asociacion Centro De Tecnologias De Interaccion Visual Y Comunicaciones-Vicomtech, Spain; I.S.I. Nederland B.V, Netherlands; Trivisio Prototyping Gmbh, Germany; Instituto de Engenharia de Sistemas e Computadores - Investigacao e Desenvolvimento em Lisboa, PORTUGAL; Graphitech, Italy; The Lighthouse, United Kingdom; Fraunhofer Gesellschaft, Germany.

© Wole Oyekoya

4. PROFI

Perceptually-relevant Retrieval of Figurative Images (such as clip art, logos, signs).

Project details

Project Reference: 511572	Contract Type: Specific Targeted Research Project
Start Date: 2005-01-01	End Date: 2007-12-31
Duration: 36 months	Project Status: Execution
Project Cost: 1.03 million euro	Project Funding: 997000.00 euro

Participant Organisations

University of York United Kingdom; Aktor Knowledge Technology NV Belgium; Freie Universitaet Berlin Germany; Universiteit Utrecht Netherlands.

5. REVEAL THIS

Retrieval of Video and Language for the Home User in an Information Society. Cross-media content for leisure and entertainment.

Project details

Project Reference: 511689	Contract Type: Specific Targeted Research Project
Start Date: 2004-11-01	End Date: 2007-04-30
Duration: 30 months	Project Status: Execution
Project Cost: 3.49 million euro	Project Funding: 2.30 million euro

Participant Organisations

Xerox - The Document Company France; Canal+ Belgique Belgium; Sail Labs Technology AG Austria; Tveyes UK Ltd United Kingdom; Katholieke Universiteit Leuven Belgium; University of Strathclyde United Kingdom; Institute for Language And Speech Processing Greece.

6. S2S^2

Sound to Sense, Sense to Sound

Nowadays, there is a wide variety of techniques tat can be used to generate and analyze sounds. The CA S2S^2 has been conceived to prepare the scientific grounds on which to build the next generation of scientific research on sound and its perceptual/cognitive reflexes. So far, a number of fast-moving sciences ranging from signal processing to experimental psycology, from acoustics to cognitive musicology, have tapped the S2S^2 arena here or there.

Project details

Project Reference: 3773	Contract Type: Coordination action
Start Date: 2004-06-01	End Date: 2007-05-31
Duration: 36 months	Project Status: Execution
Project Cost: 1.53 million euro	Project Funding: 1.30 million euro

Participant Organisations

Helsinki University of Technology Finland; Kungliga Tekniska Hoegskolan Sweden; Ecole Normale Superieure Paris France; Oesterreichische Studiengesellschaft Fuer Kybernetik Austria; Universita Degli Studi Di Padova Italy; Universita Degli Studi Di Verona Italy; Universiteit Gent Belgium; Fundacio Universitat Pompeu Fabra Spain; Universita Degli Studi Di Genova Italy; Universite De Bourgogne : Dijon France; Firenze Tecnologia ITALY.

7. IP-RACINE

Integrated Project - Research Area CINE. Cross-media content for leisure and entertainment

IP-RACINE will extend the state of the art and enhance European competitiveness by creating technologies to deliver enhanced 'cinematic' entertainment that is transferable cross-platform.

© Wole Oyekoya

Project details
Project Reference: 511316 Contract Type: Integrated Project
Project Cost: 14.36 million euro Project Funding: 8.60 million euro

Participant Organisations
Barco Nv Belgium**;** Limburgs Universitair Centrum Belgium; Deutsche Thomson-Brandt Gmbh Germany; Radamec Broadcast Systems Limited United Kingdom; Thomson Broadcast & Media Solutions Nederland B.V. Netherlands; Evs Broadcast Equipment Belgium; Filmlight Limited United Kingdom; Fondazione Scuola Di San Giorgio Italy; Pandora International Limited United Kingdom; Mediaproduccion Sl Spain; Quantel Limited United Kingdom; Joanneum Research Forschungsgesellschaft Mbh Austria; The University Of Glasgow United Kingdom; Fundacio Universitat Pompeu Fabra Spain; Universitat Pompeu Fabra Spain.

8. PENG
Personalised News content programming
The PENG project aims at defining a flexible, personalised and context-aware system for the gathering, filtering, retrieval and presentation of multimedia news for news professionals (e.g. journalists and editors), with a view of making the system also available later for general users.

Project details
Project Reference: 4597 Contract Type: Specific Targeted Research Project
Project Cost: 1.77 million euro Project Funding: 1.12 million euro

Participant Organisations
University Of Strathclyde United Kingdom**;** Atos Origin Sociedad Anonima Espanola Spain; Universite Joseph Fourier Grenoble I France; Societa Svizzera Di Radiotelevisione (Rtsi) Switzerland; Universita Della Svizzera Italiana Switzerland; Consiglio Nazionale Delle Italy.

9. SIMAC
Semantic Interaction with Music Audio Contents. Semantic-based knowledge systems
The SIMAC project addresses the study and development of components for a music information retrieval system including extensive usage of semantic descriptors of musical content.

Project details
Project Reference: 507142 Contract Type: Specific Targeted Research Project
Duration: 27 months
Project Cost: 2.98 million euro Project Funding: 1.90 million euro

Participant Organisations
Philips Electronics Nederland B.V. Netherlands; Oesterreichische Studiengesellschaft Fuer Kybernetik Austria; Matrix Data Limited United Kingdom; Queen Mary And Westfield College, University Of London United Kingdom; Fundacio Universitat Pompeu Fabra Spain.

10. AMI
Augmented Multi-party Interaction. Multimodal interfaces
AMI is concerned with new multimodal technologies to support human interaction, in the context of smart meeting rooms and remote meeting assistants.

Project details
Project Reference: 506811 Contract Type: Integrated Project
Duration: 36 months
Project Cost: 16.82 million euro Project Funding: 8.80 million euro

© Wole Oyekoya

Participant Organisations
Netherlands Organisation For Applied Scientific Research - Tno Netherlands; The University Of Sheffield United Kingdom; Fastcom Technology S.A. Switzerland; Deutsches Forschungszentrum Fuer Kuenstliche Intelligenz Gmbh Germany; Philips Electronics Nederland B.V. Netherlands; Novauris Laboratories Uk Limited United Kingdom; Spiderphone Sa Switzerland; Idiap (Fondation De L'institut Dalle Molle D'intelligence Artificielle Erceptive) Switzerland; Realvnc Limited United Kingdom; International Computer Science Institute United States; Geie Ercim France; Vysoke Uceni Technicke V Brne Czech Republic; Universiteit Twente Netherlands; Technische Universitaet Muenchen Germany; The University Of Edinburgh United Kingdom.

11. MUSCLE
Multimedia Understanding through Semantics, Computation and Learning. Semantic-based knowledge systems.
MUSCLE aims at creating and supporting a pan-European Network of Excellence to foster close collaboration between research groups in multimedia data mining on the one hand and machine learning on the other in order to make breakthrough progress.

Project details
Project Reference: 507752 Contract Type: Network of Excellence
Duration: 48 months
Project Cost: .00 euro Project Funding: 6.90 million euro

Participant Organisations
Technical Research Centre of Finland**,** Finland; Universiteit Van Amsterdam, Netherlands; University of Ulster, United Kingdom; Institut National De Recherche En Informatique Et En Automatique, France; University College London, United Kingdom; Technische Universitaet Wien, Austria; Technion - Israel Institute of Technology, Israel; Universitat Politecnica De Catalunya, Spain; The University of Surrey, United Kingdom; Consiglio Nazionale Delle Ricerche, Italy
Albert-Ludwigs-Universitaet Freiburg, Germany; Foundation for Research And Technology - Hellas, Greece; France Telecom, France; Technische Universitaet Graz, Austria; Academy of Sciences of The Czech Republic - Institute of Information Theory And Automation, Czech Republic; Aristotle University of Thessaloniki, Greece; The Chancellor, Masters And Scholars of The University of Cambridge, United Kingdom; Arc Seibersdorf Research Gmbh, Austria; Ltu Technologies, France
Association Pour La Recherche Et Le Developpement Des Methodes Et Processus Industriels, France
Ecole Nationale Superieure De L'electronique Et De Ses Applications, France; Telecommunications Systems Institute, Greece; Advanced Computer Vision Gmbh - Acv, Austria; Institut Fuer Bildverarbeitung Und Angewandte Informatik E.V., Germany; Bilkent Universitesi, Turkey; Magyar Tudomanyos Akademia Szamitastechnikai Es Automatizalasi Kutato Intezet, Hungary; Groupe Des Ecoles Des Telecommunications, France; The Provost Fellows And Scholars of The College of The Holy And Undivided Trinity of Queen Elizabeth Near Dublin, Ireland; Commissariat A L'energie Atomique, France; Centre National De La Recherche Scientifique, France; Stichting Centrum Voor Wiskunde En Informatica, Netherlands; Kungliga Tekniska Hoegskolan, Sweden; Institute of Communication And Computer Systems, Greece; Tel Aviv University, Israel; Geie Ercim, France.

12. Audio-Visual Speech Recognition in the Presence of Non-Stationary Noise
EPSRC Grant Reference: GR/T04823/01
Starts: 1 July 2004 Ends: 30 September 2006 £ Value: 116,853
Award Type: First Grant Scheme
EPSRC Research Topic Classifications: Human Communication, User Interface Technologies, Vision, Hearing and Other Senses.

Department: Computer Science
Organisation: University of Sheffield

© Wole Oyekoya

Abstract: This proposal concerns the development of novel techniques for exploiting visual speech information (e.g. lip and face movements) in the design of automatic speech recognition systems. The proposal extends this approach into the audio visual domain.

13. eSIGN – Essential Sign Language Information on Government Networks
Coordinator: University of Hamburg, Institute of German Sign Language and Communication of the Deaf

eSIGN Consortium: U Hamburg (DE); Systematics Integrations (EDS subsidiary, DE); Norfolk City Council (UK); Viataal (NL); U East Anglia (UK); Televirtual (UK); Royal National Institute for Deaf People (UK).

© Wole Oyekoya