THE OBJECTIVES AND SUMMARY FOR THE PROPOSAL ARE TO BE FOUND ON THE OFFICIAL PROPOSAL FORM.

BACKGROUND

This proposal follows from the establishment at Sheffield of the Centre for the Understanding of Digital Objects (CUDO). CUDO is a collaboration between the Department of Computer Science and a large number of highly rated departments in the Faculties of Pure Science, Architecture and Arts. Included in this collaboration is the Humanities Research Institute (HRI) and many existing external research collaborators.

Science, engineering, innovation and scholarship are increasingly dependent on computer-based representations of complex structures, whether derived by data capture from the real world, or created from scratch, or a mixture of the two. We introduce the term *digital objects* to refer to these structures. Examples of these objects are: Visual, audio and audio-visual archives; Virtual Buildings; Archaeological Objects; Textual Archives... etc. The mission of CUDO is to "understand" digital objects: their *design, creation, analysis, maintenance, manipulation* and *application*.

Studies based on digital objects require *active interaction* rather than *passive absorption*, for instance:

- searching by voice/image in audio-visual archives, resulting in adding *links* into the material, so that the material is enriched and thus more exploitable;
- manipulating and maintaining 3D designs in architecture to assist with the continual revision of aspects of the design and overcoming the many conceptual obstacles to understanding the properties of potential buildings;
- examining how created objects can be represented and manipulated so as to study the creation process itself, a major theme here will be music and sound and literary texts and artistic artefacts;
- extracting and manipulating information from digitised manuscripts.....

In this proposal we seek support for an integrated set of projects within the CUDO remit.

PROGRAMME AND METHODOLOGY

The projects will be located in CUDO, which will have dedicated SRIF-funded facilities in the University's Multi-disciplinary Research Centre, housed in refurbished premises previously occupied by the research laboratories of the Health and Safety Executive. Included in these premises will be a state of the art immersive virtual reality facility, (CAVE), a sophisticated sound studio and dead room, other digitisation facilities and laboratories for all the researchers. There will be access to broadband computing facilities, in particular the new White Rose Universities supercomputers and GRID. The portfolio of projects include 2 *core technology* activities (*A* and *B*) together with a number of individual application-driven projects in specific subjects and two *integrator* projects (*I* and J) which will develop full multi media systems based on the three key media of image, text and sound. The themes addressed in the *Project Summary* are indicated.

TECHNOLOGICAL FOUNDATIONS.

A. Computer Science - *Digital objects: standards, representations and repositories. [Holcombe et al].* {1 RA + 1 PhD}

The development of Digital Objects technology and Multi Media generally is progressing at an exceptional rate, stimulated by the wide availability of powerful desktop computers and the internet. It is vital that the emerging standards such as MPEG-7, MPEG-21 ..., languages such as SMIL (Synchronized Multimedia Integration Language) and software to support these are involved at the outset if the work is to be *future proofed*. We intend to develop repositories of digital objects using new technologies based on XML. A key issue will be to develop these repositories to support the research into the application of digital objects in the other projects. Thus the role of the research fellow, together with the specialist computer science researchers, will be to identify, develop and transfer these technologies into the project teams. The repositories will include:

Digital Sound. Sheffield has a world class reputation [Computer Science] for building robust information *retrieval* systems for speech archives (with BBC collaboration) which operate in real time and provide querying from primary sound sources. The project will exploit and extend this technology.

Digital Text archives. Sheffield is also a world-leader [Computer Science] in information *retrieval* from text. It is a major partner in the EPSRC IRC *AKT* project. The project will exploit this technology. In addition, the HRI has world leading specialisms in working with non-standard orthographies, in creating multi-faceted interfaces for working on the genetics of literary texts, on collation routines, etc.

Digital 3d objects archives. Sheffield has a world class reputation [Architecture, History, HRI] for constructing virtual reality models of built and historic environments. These require sophisticated repositories of 3D images. The project will exploit and extend this technology.

Further work is needed on evaluating recent advances in content based image retrieval (CBIR) and this will include feasibility studies involving existing WWW systems such as WebSeek, Photobook, etc. based on colour, shape and texture criteria. Further work is also required on XML and its derivatives and supporting standards and software to provide smart methods for exploiting the use of digitised resources and archives. We have built XML-based repositories of digital objects and have developed methods for the rapid evolution of such repositories to address the changing needs of the application. The Description Definition Language for MPEG-7, W3C's Resource Description Framework and the MPEG-21 Multi Media Resource Framework will be important drivers for the implementations developed.

Successful techniques from Extreme Programming, piloted at Sheffield, [Holcombe] will be encouraged in much of the software development activities. The quality of academic software, whether produced by computer science researchers or by researchers in other fields is often of poor quality - unstable, hard to use and difficult to exploit by non-specialists. Partly this is due to the informal nature of the way it is developed. Extreme programming is a new, rigorous but highly agile software engineering technique that emphasises robust testing, close client contact and incremental delivery. It is an ideal method for the engineering of this sort of application driven software.

We will build a virtual guide for a digital object collection, an interface involving a virtual actor who would guide the user

around the collection. This will involve a talking head to converse with the user. It could potentially be used as a front-end in conjunction with the other retrieval-based work, or as a front-end guide in some of the other projects. **Themes 1, 2, 5**.

B. Computer Science *- Data-driven content based retrieval. [Niranjan et al]*{1 RA + 1 RA linked with I (Comp. Sci.)}
This project will address techniques for the efficient organisation of complex digital objects for rapid, content addressable retrieval based on partial or uncertain cues. An example would be the search through video archives for clips containing Clinton and Lewinski, which is reported to have taken CNN over 10 days to perform manually. With multi modal digital objects containing speech, music and images, searching with partial information can be efficient if the data is organised in a hierarchical structure reflecting its inherent statistical regularity. An example of such regularity is the concept of "eigen faces" to represent facial images. Recent research has shown how large collections of facial images can be represented accurately via small numbers of orthogonal face-like images. These representations have been used effectively in the classification of faces as well as low bit rate representation of their images.
In this core project we will explore machine learning techniques that are suitable for extracting complex statistical regularities from the sort of repositories of multi modal digital objects being developed in other projects. We will pursue techniques for defining distance metrics that reflect statistical regularities by representations that are based on probabilistic generative models of the data. Retrieval with partial cues can then be cast as a probabilistic inference problem involving the imputation of missing data. We will implement such a framework centred around the machine learning technique of *Kernel-machines* which combines probabilistic inference and powerful results from nonlinear function approximation. Digital objects consisting of video, text and sound will be used as test beds to develop algorithms and critically evaluate them in the light of their use by the other applications. **Themes 1, 2, 5.**

These strengths and achievements will be built upon and applied in the individual projects and combined in the integrator projects to create full multi media systems.

Specific applications projects.

We envisage a number of projects utilising specific types of digital technology as well as two integrator projects which will involve images, text and sound.

C. Archaeology *- Digital archiving, modelling and visualisation of artefacts and sites. [Dennell et al]*.{1RA}
We are concerned here with digital archiving of historical and archaeological research resources and with digital research resources in medical history. The project will contribute to the development of a *digital archive and on-line database centred on the design, manufacture and technological improvement of early modern surgical instruments*, incorporating materials from the Hawley Collection and other curatorial institutions.

Computer-Aided visualisation of archaeological Sites and the development and application of digital imaging technology for detecting, mapping and representing complex interior spaces in archaeological sites. The project will broaden existing research on archaeological caves to incorporate prehistoric megalithic monuments and other built structures of archaeological importance. Collaborators: Dr. William Sellers, Department of Human Sciences, Loughborough University.

Modelling of the impact of construction activity on archaeological resources and the characterisation of the impact of construction activity on archaeological sites and artefacts. The project will centre on the manipulation of images from engineering model studies to assess the deformation of archaeological sediments through constructional piling, and to assess damage to artefacts as a result of sediment loading. Collaborators: Dr. Adrian Hyde, Geotechnics Research Group, Department of Civil and Structural Engineering.

Archaeological and forensic facial reconstruction and quantitative comparison of facial profiles from skulls. Existing collaborative research on computerised forensic facial reconstruction will provide the starting point for novel approaches to the quantitative comparison of facial profiles scanned in three dimensions from archaeological skeletons. Collaborator: Dr. Martin Evison, Department of Forensic Pathology, University of Sheffield. **Themes 1, 2, 3, 4, 5.**

D. Music *- composition and performance [Clarke et al].* {1 3year RA jointly with Computer Science + 1 PhD}
(i) Data retrieval from archives of recorded musical performance.
The study of the history and aesthetics of performance has been transformed by the development of the recording industry, and by the presence of substantial archives of recording in many parts of the world that document a hundred years or more of musical performance. These archives are increasingly based on digital formats of one kind or another which, together with digital editing systems, offer the prospect of powerful methods for documenting and analysing the history of musical performance. However, the detailed analysis of musical performance from acoustical sources is still in its infancy, and current methods are haphazard, diverse, lacking in power and sophistication, and extremely labour intensive. The aim of this project, in close conjunction with research in the recognition of speech and gestures, is to develop more effective, labour-saving and powerful methods for retrieving data (primarily in the domains of timing, intensity, and timbre) from digital encodings of recorded performances. These have the potential to make a major contribution to an understanding of musical performance. Joint with Computer science. **Themes 1, 2, 3, 4, 5.**
(ii) Modelling virtual instruments and acoustical environments
From its origins in pure tape composition, electroacoustic music has developed in a number of important directions. A significant factor here is the relationship between electroacoustic music's more abstract computer-based implementation and the character and influence of instruments and acoustical environments. Electroacoustic music has increasingly involved the use of virtual instruments as compositional tools and the interaction between computers and live instruments. In a similar manner, virtual acoustic environments (which have been increasingly exploited by studio techniques across a

whole range of musical genres) and the real acoustical spaces in which electroacoustic music is diffused, have become increasingly important as a focus for research and creative work. This research will take forward research in virtual instruments and virtual environments with the emphasis on the creative possibilities that virtual modelling can afford. Link to project *I*. **Themes 1, 2, 3, 4, 5.**

(iii) Creative manipulation of digital sound objects

The theories behind the creative manipulation of digital sound objects (*l'objet sonore* as coined by Pierre Schaeffer) have lost momentum over recent years. Schaeffer's treatise still only exists in French and very few practitioners and theorists have broadened its horizons. The quest for theories of expression within domains that have very few boundaries is one that should go hand in hand with some form of practice. The research strand here seeks to explore the framework for an electroacoustic amanuensis, whose purpose would be to monitor and prompt the computer-based composer and provide shortcuts for commonly used compositional functions. **Themes 1, 2, 3, 4, 5.**

E. English - *Relatedness amongst languages. [McMahon et al].* {1RA + 1 PhD - one joint with Comp. Sci.}

Traditional representations of relatedness among languages typically involve the classical, two-dimensional family tree. Trees indicate common inheritance of features, but cannot show the results of language contact; and they are still usually hand-constructed by experts in the structure and history of the language family concerned. There is an increasing need to automate this process, to make it applicable to language groups without a significant written history, where relationships are unclear. A current project is already interdisciplinary in unifying linguistic data with tree-drawing and tree-evaluation programs from biology, and over the next two years the project team aims to expand this work to focus on networks as well as trees, and on cluster analysis. However, these representations remain two-dimensional, and prioritise either common ancestry or contact phenomena. This project involves collaboration with a computational linguist to explore three-dimensional (or multi-dimensional) modelling with either a tree or network basis. Such three-dimensional digital objects might allow inherited and borrowed forms to be shown together, and to be quantified, leading to considerable advances in our understanding of the relationships between languages. More complex models of this kind would also allow handling of more complex data, meaning that the radical summarising required in manual tree-drawing, and the usual concentration on word-lists, can be replaced by a fuller consideration of real and representative data on words, sounds, and grammar. **Themes 1, 2, 3, 4, 5.**

F. Humanities Research Institute - {*1 3year RA + 1 PhD + 1 PhD linked to Architecture}*

The Humanities Research Institute is a University Research Centre devoted to the creation of advanced electronic editions of text and image-based corpora in the humanities that will shape the virtual cultural inheritance of the future. The four projects outlined below enable established projects to undertake research that will exploit to the maximum electronic materials already in existence or in creation. The use of XML is vital and there will be a strong emphasis on the development of 'generic tools' for handling digital objects. The RA will be a specialist in historical virtual reality.

(i) The Cistercians in Yorkshire: a virtual monastic environment. [Foot et al.].

This project investigates the intellectual contribution that three-dimensional virtual-reality models can make to our understanding of the architectural history of an inter-related group of now ruined monasteries in Yorkshire, by setting the scholarly reconstructions in a broad textual and artefactual context. These reconstructions will be created from site plans, antiquarian drawings and early photographs as well as the surviving standing (and fallen) fabric. Different models will illustrate the separate phases of architectural development of the ecclesiastical and claustral buildings between their twelfth-century foundation and the Dissolution of the Monasteries, thereby contributing to scholarly debate about the architectural history of each monastery. A digital archive for the Cistercians in Yorkshire will be created, this will seek to digitise a range of texts, visual images and artefacts illustrative of Cistercian thought, culture, economy and daily life, illustrations in Cistercian manuscripts from the region and beyond and archaeological and art historical artefacts. This project is genuinely multi-faceted and interdisciplinary. **Themes 1, 2, 3, 4, 5.**

(ii). Imaging and identification of masonic and friendly society artefacts. [Prescott et al] - The Centre for Research into Freemasonry in the Humanities Research Institute has special access to the collections of the United Grand Lodge of England, the most important collection of masonic books and artefacts in the world. Friendly societies are of increasing interest to social historians because of their importance as a means of working-class self-organisation. The Centre for Research into Freemasonry will seek to use DOSITS resources to investigate 3D imaging of regalia and jewels, and to examine how far automatic shape recognition techniques can be used to assist in the sorting and identification of large quantities of material bearing symbols of this kind. **Themes 1, 2, 3, 4, 5.**

(iii). Digital Representation of Object Marking. [Greengrass & Luscombe] - Unlike the London cutlers, Sheffield's Company of Cutlers have retained the markbooks of their cutlers from the incorporation of the company in the early seventeenth century. These mark-books form a unique record in England of the identification framework utilised in the cutlers' trades for the range of products produced by cutlers in the county of Hallamshire. This project will use fractal analysis to describe the range of cutlers' marks as a prelude to developing a system of understanding of such marks that would be of wider application to silver hallmarks and furniture marking. The project is currently based in the Hawley Research Centre in the Archaeology Department but with links to the Humanities Research Institute. **Themes 1, 2, 3, 4, 5.**

(iv). Mapping 18th Century Sheffield. [Greengrass & Luscombe] The Fairbank Papers contain the archives of the firm of Fairbank, land surveyors from 1737-1848. This is a unique surviving collection of maps, account books, correspondence and apprentices' work diaries as well as field notebooks from the late eighteenth and early nineteenth centuries. Using the catalogue and calendar that has been created from the 360 field books surviving, this project will use a combination of CAD, digitised recreations of techniques used by surveyors of the period, present-day site photograph etc to research the processes, and evaluate the accuracy, of mapping techniques of the period. **Themes 1, 2, 3, 4, 5.**

G. German - *Translating German texts between multiple text and script forms. [Perraudin, Newton].* {1 RA}
In Germany, many older literary and other printed works are effectively inaccessible to non-scholarly contemporary readerships because they only exist in *Fraktur*, the 'Gothic' script which finally disappeared from general use after 1945. In English and French, out-of-copyright texts which are uneconomic to retype but are of interest to specialist modern markets can readily be reissued by photo-reproduction; in Germany this is not possible, as readers will not tolerate *Fraktur.* The modest attempts so far undertaken to make text in *Fraktur* digitally readable, and thus immediately printable in Roman script, have been fairly unsuccessful. This remains a major barrier to reproducing material conveniently. If the latest text-reading software can be made trainable to cope with *Fraktur* typefaces (and with the associated, slightly deviant orthographies), then original works will become accessible again. Particularly of relevance are non-canonical literary and semi-literary prose publications of the 19th century, such as women's novels, proletarian literature, travelogue and memoir, which were long ignored but in which modern readerships are becoming interested, in line with new developments in taste. The proposed project aims to compare the utility of different text-reading software for reading *Fraktur* and to devise ways of training the best such software for specific variants of the form, so that substantial texts can be converted quickly and precisely. It will use as its textual material, initially, printed works in Sheffield University Library, then either works held in the British Library or (preferably) the microfiche collection *Bibliothek der deutschen Literatur,* which incorporates all German literary publications from 1750 to 1900. Objectives of the project are the broad cultural aspiration of extending awareness among modern German readerships of earlier, now-neglected creative writing, and, more practically, a plan to develop a text-transcription service available to German publishers. The two proposers are experts in 19th century German literature and society, with a focus on non-canonical literary forms and their publishing contexts (Perraudin), and in German philology, with a particular interest in the history and morphology of *Fraktur* (Newton). **Themes 1, 2, 3, 4, 5.**

H. Biblical studies*. Analysing and exploring biblical documents and paintings.* {1 PhD}
(i) *Biblical paintings and music [Exum].*
There are many historical paintings of biblical personalities and events in museums and other archives. These paintings contain a wealth of important information about the context and cultural background to biblical interpretation past and present. Applying a range of theoretical models to interpreting the Bible, including gender criticism, reader-response, narratology, psychoanalytic criticism, queer theory, film, and visual art theories, the project proposes to digitise such paintings and to analyse and annotate them to provide an important resource for biblical scholars. The relationship with ancient church and monastic music will also be explored to provide a full multimedia resource based on the paintings, textual critical studies, music texts, libretti and liturgy, together with sound recordings and annotations. The aim is to create a catalogue of works of fine art in British provincial art galleries with biblical themes, with links to musical adaptations of the particular texts represented, for the use of researchers in the newly developing field of the influence of the Bible on culture. The context of the research is exploration of the Bible as a cultural icon, considering both its impact upon culture and the effect of cultural conventions upon its interpretation. The availability of this material as a resource will generate a number of significant studies of interest both to biblical scholars and to art historians, from gallery materials that have been seriously underexploited in the past. It is expected that the project will uncover ample materials for comparative and analytical study by advanced scholars who will appreciate the addition of works on biblical themes by lesser-known artists to the already sizeable collection of biblical works by the most eminent of painters. There are clear links with projects such as the Cistercians in Yorkshire and the aim is to promote joint investigations into these and to create powerful environments for greater understanding of religious life in this and other periods. **Themes 1, 2, 3, 4, 5**.

(ii). Digitizing Ancient Syriac Manuscripts of the Biblical Book of Judith [Clines].
Clines is a member of an international team of scholars preparing a critical edition of the ancient Syriac Bible, which is known as the Peshitta version. This critical edition reports all the variants found in the manuscripts of this translation of the Bible. Some twenty volumes of the edition have already been published. For the Book of Judith, one of the books of the Old Testament Apocrypha, more than 50 manuscripts survive. Many of the manuscripts are in excellent condition, and written in the very clear Estrangelo script of Syriac. The first aim of the present project would be to create a machine-readable text of each of these manuscripts by developing optical character recognition software appropriate for Syriac and capable of handling the right-to-left orientation of its Semitic script. The second aim of the project would be to adapt current software for comparing variant texts to the special needs of this script, so that the largely mechanical work of comparison and recording of variants can be carried out more accurately. The complete text of all the manuscripts would also be available to other scholars both as images and as digitized texts; this would be a scholarly advance on the present situation, where scholars are dependent for statements of the textual evidence upon the eyesight and accuracy of the individual editors of the biblical books, and have no way of checking their results. **Themes 1, 2, 3, 4, 5.**

I. *Integrator: Architecture -* *MOVE - Multiplex Open Virtual Environment for architecture.[Lawson et al].* {1 RA + 3 PhDs (some linked to Comp. Sci.) + 1 PhD linked to History}

There will be three related strands of work, the first being the development of visual and acoustic models of buildings and the urban environment and the validation of these models against real-world human experience. In the second strand, we shall further explore visual interactive models of buildings particularly studying the way in which they can be captured with low levels of labour from the real world or from drawings. Lastly, we will explore how such models can be established from many locations with many contributors and how such models can be made available interactively across the Internet. Tapping into the new Web standards (e.g., XML, SVG, X3D, MPEG-4) and Internet-enabled open platforms, the visual modelling research aims to develop the basic technology for creating a Multiplex Open Virtual Environment (MOVE). MOVE will extend conventional Virtual Environment (VE) capabilities to allow simultaneous multi-dimen-

sional modelling (e.g., 2D mapping, 3D modelling, 3D modelling in time with ambient intelligence transmitted in images, sounds and texts). With MOVE, greater end-user control, semantic processing, and data reusability can be achieved. Interacting with an instance of MOVE, end-users will be able to reconstruct parts of the virtual world through on-line reconfiguration of different datasets retrieved from distributed data repository servers. The MOVE technology will facilitate multidisciplinary collaborative development of both open and proprietary data repositories for creating dynamic VE that can be reused for unforeseen purposes.

We shall explore the acoustic environment of both interiors and exterior urban space. We shall explore how the visual models can be used to input the geometric and material information into the acoustic models. We shall explore how the acoustic models can be represented back to users in conjunction with the visual models. That is to say users may be able to experience both visual and acoustic information about places simultaneously including urban streetscape and soundscapes. Unlike conventional acoustic modelling, we will pay particular attention to the acoustic design of a space sequence, not only a single space. In this respect, real-time auralization would be very important—this is also useful when designers want to change certain building components and then listen to the difference in acoustic performance. To achieve this, research would be needed to simplify the algorithms and to verify the simplifications.

We shall conduct research to validate these models at two levels: the *digital* level and the *human experiential* level. Models must be accurate in terms of their pure digital information. In the case of visual models the captured and represented data must match the measured data of the real world scene. In the case of acoustic models, those parameters of the acoustic environment represented will be validated against real world measurements. We will also use the techniques of acoustic physical scale modelling to validate the computer/theoretical models—in scale models configurations can be changed easily. However this still does not guarantee the digital models will accurately match the human experience of the real world. We shall perform a series of validating experiments to measure human responses in both real world and digital models to establish the validity of the digital models. This work will also extend to identifying the limits and parameters of accuracy required for effective perception. In many cases it is likely that there is a law of diminishing returns in terms of total digital accuracy. It is also essential to justify the simplifications in real-time acoustic models according to human responses. We shall establish the trade-off benefits of each type of approximation in terms of perceptual experience. Such work should enable us to arrive at the most cost effective models in terms of the work required to capture them, build them and the computing space and power needed to represent and store them.

Links between the above work with other groups. With acoustic simulation software, the music performance in different types of spaces can be created and evaluated. Acoustic facilities can be used by 3-4 groups. The basic technology for creating MOVE will have a strong link with searching by voice/image in audio-visual archives, exploring dynamic links into the material. MOVE can be exploited in archaeological and historical research and reconstruction modelling and there will be a joint researcher shared with HRI. **Themes 1, 2, 3, 4, 5.**

J. Integrator: Computer Science - *Structuring and Browsing of Meetings [Renals et al].* {1 RA + 1 PhD + 1 PhD with Comp. Sci. B}

This integrator project is concerned with the construction of a demonstration system to enable structuring, browsing and querying of an archive of automatically analysed meetings. The meetings will be recorded in a "smart meeting room" equipped with microphones and video cameras. For each meeting, audio, video, textual, and (possibly) interaction information will be available. Audio information will come from close talking and distant microphones, as well as binaural recordings. Video information will come from multiple cameras. While the video and audio information will form several streams of data generated during the meeting, the textual information - the agenda, discussion papers, text of slides - will be pre-generated and can be used to guide the automatic structuring of the meeting. The interaction stream consists of any information that can help in analysing events within the meeting, for example, mouse tracking from a PC-based presentation or laser pointing information.

To illustrate the proposed system, consider a meeting of 3-15 people, with a predefined high-level structure (such as an agenda). The browser will follow this structure: selecting an agenda item will enable a set of ways to view the segment of the meeting related to that topic. These might include a textual summary, a diagrammatic discussion flow indicating which participants were involved, or a set of audio or video key frames to give the essence of the discussion. It would also be possible to query the rest of the archive either by example from that segment, or through an explicit query.

Smart Meeting Room:

This will involve the specification and development of the meeting room (number and placement of microphones, cameras, etc.), development of the data collection protocol and annotation of the multimodal meetings database. This work will build on an existing collaboration with the International Computer Science Institute (University of California at Berkeley).

Analysis and processing of the audio and video streams:

The raw audio and video streams will be processed by multimodal recognizers including robust conversational speech recognition (to produce a word-level transcription), multimodal person identification, and source localization and tracking. The Department of Computer Science has existing expertise in these areas. Although these technologies are imperfectly developed, they are established firmly enough to warrant their use in combination.

Speech recognition is central to the proposed application, and the further development of the state-of-the-art in robust conversational speech recognition is still required. However, a multimodal approach is necessary because a large amount of information does not appear in a transcribed speech stream (eg non-speech audio or facial information). Additionally, speech in a meeting situation is characterised by an informal, conversational style: an initial word error rate (WER) of

40% (or worse) would not be unexpected in this scenario. Although some information access tasks have been accomplished using high WER transcriptions (eg, spoken document retrieval), it would be highly beneficial if information streams of other modalities were used to compensate for speech recognition errors.
Our approaches to processing these streams have some common features, including the statistical inference of recognition models (typically hidden Markov models) from data, the use of multiple streams of features (derived from multimodal sensors), and the application of feature selection algorithms. The output of these recognition processes will include "higher level" streams and annotations. Examples of higher level streams include spatial locations of sources, person identification annotations, and multiple streams of transcribed speech. The higher level streams give some information that can be used directly in the meeting browser (eg identification of participants); however integration of these streams and further processing is required to obtain much of the content of an effective browser.

Integration and structuring:
This work will involve the specification of a flexible, intelligent information management framework and the development of models for the integration of multimodal streams.
The meetings browser needs to be able to integrate information from streams of various modalities, and to enable different ways of accessing information. Beyond novel techniques for data fusion, truly multimodal approaches will require the development of new multi-channel processing approaches able to deal with multiple, non-stationary, non-synchronous, data streams. In multimodal communication it is often the case that the semantic information is spread across different modes. An objective of the integrator is the development of effective language processing techniques for utterances spread over multiple modalities, resulting in a set of partial and unreliable streams. The development of new multimodal parsing and understanding approaches, along with multisource decoding algorithms, will thus be an important objective of the present project. The construction of a system for the management of a set of parallel multimodal streams and annotations is a software problem that we address. Such an information management framework provides a basis for multimodal information access, such as textual and multimodal summarization, key phrase extraction, key frame extraction, topic segmentation and filtering, and ad hoc search across an archive. The manager will support access to the meeting recordings, through browsing time-coded summaries.

Demonstration System:
The final demonstration system will be for the offline processing of meetings. Meetings associated with CUDO will be used as a vehicle for this project. The demonstration system will make use of the advances in multimodal recognition and integration, as well our existing expertise in (and software for) information access. The software issues involved in this demonstration system will involve the use of standards such as MPEG-7 and SMIL, addressed elsewhere in this proposal.
Themes 1, 2, 3, 4, 5.

Added value component to the portfolio of projects.
It is envisaged that a number of advanced Computer Science undergraduate and masters students will be doing intensive software development projects working with researchers from all the academic areas involved in the Centre.

Relevance to beneficiaries

The potential beneficiaries can be identified as: academic researchers in the Humanities; researchers in Computer Science and those interested and involved in the exploitation and utilisation of archival content and resources based on a variety of cultural activities. Humanities researchers will benefit from having more sophisticated software support to enable them to interact in richer and more fruitful ways with their cultural material and archives. Computer Scientists will benefit from being involved in the pioneering of new multi-media systems, new languages and from interesting new challenges to the field generated from novel application ideas. Consumers and commercial organisations will benefit from increased access to rich multi media based systems containing rare and important artefacts and cultural material.

Dissemination and exploitation

The research will be disseminated through academic publications, putting both resources and public domain software on the Web. Postgraduate research and masters students will receive training during their activities in the Centre. The Government, through the regional development agencies, Yorkshire Forward and Sheffield First have identified the current high tech. services strengths of Sheffield to be in the domains of multi-media and cultural content and are preparing to invest up to £40m of EU Objective 1 in the support of expanding these activities (letter attached). The University of Sheffield has also received Objective 1 funding for supporting *spin off* companies and will be using part of the new HSL site, which will also house the Centre, for incubator units. There is, therefore, a very clear opportunity to develop commercial activities based on this research.

Justification of resources

Funds are requested for 10 RAs and 9 PhD studentships, these will be a combination of computer scientists and application domain experts. The basic *modus operandi* is that many projects will involve at least two researchers, one with a computer science background and one from an application domain to ensure that the best technology is used for the application and that there is a corresponding *pull* from the application, thus refreshing and challenging the software research area. In some cases researchers will be allocated to work with 2 projects to provide deeper integration of the work. Sharing the same premises, in CUDO, will also provide integration. Because of the complexity of the proposal and the large number of collaborators funds are also requested for an *Administrator* (50%) and a *Computer officer* (50%). Funds are also requested for some key computing technology, a large fileserver and data store together with desktop computers for all the researchers. Consumables and travel complete the request.