Roles and Reusability of Video Data in Social Studies of Interaction

SCARP Case Study No. 5

Angus Whyte

Digital Curation Centre, University of Edinburgh

DCC SCARP CASE STUDY REPORT

Deliverable B4.8.2.2

Version No. 1.1

Date 19 October 2009

Copyright

Catalogue Entry

Title Roles and Reusability of Video Data in Social Studies of Interaction

Creator Angus Whyte (author)
Subject Data curation; formats, processes and issues; system development; standards; legal factors; methodology, and problems overcome; human factors
Description This case study has been produced for the Digital Curation Centre (DCC) SCARP project, funded by the Joint Information Systems Committee (JISC) to investigate disciplinary attitudes and approaches to data deposit. The study reviews the curation landscape in several interdisciplinary fields that use video analysis in studies of human interaction. Field research for the study involved semi-structured interviews with researchers and participant observation in their informal data sharing and analysis sessions. Researchers were working in a variety of projects across social sciences, humanities and computing; from domains including human geography, psychology, sociology, social informatics, and linguistics. The study primarily focuses on uses of video in ethnographic studies and in eye movement research.

.
Date July 2009 (creation)
Type Text
Format Adobe Portable Document Format v.1.3
Resource Identifier ISSN 1759-586X
Language English
Rights © 2009 DCC, University of Edinburgh

Citation Guidelines

Whyte, A. (2009), " Roles and Reusability of Video Data in Social Studies of Interaction. SCARP Case Study No. 5", Digital Curation Centre, Retrieved <date>, from http://www.dcc.ac.uk/scarp

Contents

Executive Summary

Social science researchers are making increasing use of digital video. All of us, researchers or not, have an alluring range of commercial web sites for sharing video, although these do not cater for long-term reuse of video in research. But what kind of roles does video fulfil as research data? And what curation issues and challenges does video raise for researchers and their institutions? The phenomenal growth in public use of digital video is a topic of social research; in the first six months of 2008, users of Youtube uploaded more video footage than the top three U.S. TV networks would have broadcast if they had been operating 24 hours per day over their sixty-year lifespan (Wesch, 2008). Yet there have been few studies of social scientists' own uses of digital video data in their research.

Video data is rich in potential for repeated study within and between projects. Aiming to contribute to developing that potential, the report explores several fields of social research where video corpora are being developed, and possibilities for secondary analysis are being actively explored. Despite recent e-Social Science work in tools for analysis of multi-modal corpora, previous studies of qualitative data reuse in social science have given little attention to this area.

An increasing range of social science and technology related fields are using corpus-based approaches[1], describing and analysing patterns in and across examples of human activity, recorded in text, sound, still and moving image, and as traces of digital interaction. Video is at the core of multimodal corpora, a backing track or temporal map, allowing different views and time-based data sources to be overlaid; facilitating inter-disciplinary reuse.

The growing inter-disciplinary use, complexity and size of video data make it important for research data services to understand and support it. The report uses the DCC Curation Lifecycle Model to identify shared needs for support with curation. It also highlights disciplinary differences within the relatively small area of interaction research, pointing to diversity in the roles of video and the contextual information needed to reuse it.

The study focuses on two main groups of researchers; SEDIT (Scottish Ethnomethodology, Discourse and Interaction) is an informal cross-disciplinary network of researchers, some of whom use video as observational data for ethnographic analysis in Human Geography and in Computing. The second group use video as experimental data in studies of eye movements and scene perception, based in the Visual Cognition Research Group of the University of Edinburgh Dept of Psychology.

Chapter 1 profiles the researchers, projects and groups that were the focus of the study. Data for the case study was obtained from interviews with members of both groups, and from participant observation in SEDIT 'data sessions' where video data are collaboratively analysed. The chapter includes a literature and landscape review of developments in the fields mainly concerned, providing background on the relevance of video and the rationale for sharing data. Among these drivers for video data curation, the chapter describes e-Social Science projects addressing researchers' needs for real-time data sharing and analysis tools.

Chapter 2 presents three main themes that arose from the case study. Interviews and observations were used to identify challenges that video data poses for a 'lifecycle management' approach to planning data curation. The first theme is the *diversity of research practices* involved. Video serves different roles across and within research fields, and at different stages of a project. It may be publicly accessed and used at the beginning of its life as data (e.g. as web video clips), or made public at the end of the lifecycle (e.g. clips on researchers' websites), and shared at various points in-between (e.g. at 'data sessions' and conferences). Researchers with broadly similar analytic orientations, e.g. to ethnographic

[1] This definition is meant to include any methodology where collected 'examples of human activity' are produced as a resource for repeated analysis.

observation, use video material differently according to the research topic or application area. Technology choices were more closely related to researchers' communities of interest or practice than to wider disciplinary contours; especially as researchers import methodologies from disciplines other than that which their research group is institutionally aligned with.

The second theme is the *uncertainty* affecting data curation planning decisions based on the Curation Lifecycle model, given rapid changes in technology, the complexity of the format choices to be made, and the need in exploratory research to begin with open questions about the data to be acquired and analysed. Methodological preferences and research topics influence the level of image detail needed to describe, analyse and interpret video alongside other data associated with it, depending for example on the required attention to the detail of gestures, or shifts in gaze.

Methodology similarly affects *appraisal and selection* of video data. In ethnographic studies data quality judgements involve trade-offs between the unfolding relevance of the material, the audio and visual clarity, and legal or ethical factors. Similar factors affect experimental psychology where video is used as an experimental stimulus. Here the need to maximise image clarity is driven by the need to perform statistical analysis on frame-by-frame changes in the image content, correlating these with other experimental variables. Studies in this field are 'data driven', using hypotheses based on exploring the patterns found rather than on theoretical models. Finding moving images to use as experimental stimuli that are controllable, ecologically valid, and are likely to yield informative results involves trial and error with a variety of content genres. Meanwhile the eye movement research community is only beginning to formulate expectations of how results should be made available.

Researchers generally did not plan for long-term *preservation* given the uncertainties of post-project funding and the confidentiality, consent and copyright issues in sharing video openly. In both observational and experimental studies, how much of acquired video data can be shared beyond the research team is a legal/ethical question as much as it is about the infrastructure for sharing video. Storage constraints affect all parts of the curation lifecycle for video data, each step entailing questions about 'where it will fit now' given the available and affordable capacity, and 'where it could go later' given the ethical limitations on disclosure.

The *third* theme of Chapter 2 is the nature of the *context information* needed to reuse archived video objects. Much social research that uses video is concerned with understanding phenomena in their natural setting. This makes it problematic to differentiate between data and context, especially if context is treated as a static description. Social research has different perspectives on what 'context' is, and whether it can ever be satisfactorily recorded. Video archives would be better enabled to address this by treating context information as a dynamic property, resulting from dialogue between the original researcher and reusers.

Chapter 3 considers the implications for curation lifecycle managing. To manage the range of possibilities and contingencies, a more iterative approach is proposed involving three main cycles of curation. Firstly a '*planning and piloting'* phase begins with the data management plan and then revises this in light of the data initially gathered. The main '*project curation*' phase begins with selection of data for analysis and implementation of tools and standards to enable involvement of colleagues and peers in that. Then the '*long term curation*' phase begins as researchers' work up the data for publication and longer-term preservation and reuse.

The report summarises curation strategies researchers in this study adopt, and provides sources of further guidance, drawing on the literature including recent landscape reviews for JISC and the AHRC. It also draws on discussions with University of Edinburgh research data service providers aiming to envisage how, in this and other UK institutions, the respective roles of research groups and centrally provided data archives may evolve to handle video material.

Conclusions and Recommendations

The richness of digital video data for repeatedly analysing human interaction is driving the development of shared data resources and tools in social research fields that are concerned with closely analysing language and interaction. These include a range of 'data driven' traditions that are finding novel ways to identify patterns in their data for further interpretation or experimentation. Just as technology has underpinned the development of corpus-based linguistics, development support for online corpora of video and related materials is likely to promote reuse of data in interaction-oriented social science, building on methodological traditions of reusing and sharing examples of interaction.

The curation needs of researchers in multimodal interaction differ in important ways from those of qualitative or mixed-method researchers in other areas. One of the main differences is in the contextual information requirements that would support reuse. In many areas of qualitative social research lack of access to the original research context is commonly seen as a major barrier to secondary analysis. In interaction analysis it is less of a barrier. Where the analysis of particular actions or behaviour depends on understanding their place in an unfolding sequence of interaction that has been audio-visually recorded, the data 'content' is itself part of 'the context'. Sharing this data depends partly on documented detail provided up-front by the data creator, but mostly on the possibilities that collaboration with others affords for developing a richer analysis of it.

The needs for collaboration support have begun to be addressed through ESRC funded work by the UK Data Archive, and in the UK e-Social Science programme. Meanwhile more advanced models for archiving and curating annotated video or multimodal[2] corpora are being developed and implemented by linguistic archives in the US, France and the Netherlands. These provide searchable corpora comprising video and synchronised data that may be browed with their annotations online, to aid and stimulate reuse. UK researchers, for example in the *DreSS*[3], and *AMI* [4]projects, are already adopting techniques to develop support for cross-disciplinary interaction analysis. Archival and metadata models from the linguistic community may also have wider influence on reuse in social interaction research.

Sustaining the accessibility and reusability of digital video-based research materials is a challenge to domain-based archiving initiatives and to national and institutional data services. The challenges include identifying and fulfilling the various roles that video may play at different stages in the research process, and enabling appropriate legal and ethical controls on data access.

Recommendation 1- DCC, JISC and other research funders should develop the e-infrastructure for multi-disciplinary interaction research by facilitating workshops to bring together the disciplines involved, disseminate relevant tools, and explore more effective ways to browse and annotate multimodal data in data repositories. This would take forward work piloted by UKDA and by the DreSS project.

When video constitutes a significant proportion of the data to be created or collected for research purposes, decisions on how to manage it are likely to be revisited repeatedly, long before any of it is archived for potential reuse. Early decisions on the data to be collected; options and formats for capture and storage, and the tools and resources for analysis are likely to change throughout a project. Methodology may require initial questions to be refined in light of patterns identified from early data collection, which may also narrow technical options. A phased approach to assessing the risks to re-usability is needed especially given

---

[2] The term 'multimodal' is favoured by linguistic and psychology researchers over 'multimedia' as it better reflects their view of audio-visual and instrumented data as recordings of 'modes of communication'.

[3] Digital Records for e-Social Science (DreSS) available at: http://web.mac.com/andy.crabtree/NCeSS_Digital_Records_Node/Welcome.html (August 2009)

[4] Augmented Multiparty Interaction (AMI) available at: http://corpus.amiproject.org/ (August, 2009)

the changing and complex relationships between data policies, ethics and rights issues.

Recommendation 2 –The DCC Curation Lifecycle Model is an 'ideal type' and rather than used as a one-off framework for Data Management Planning it should be used iteratively during research projects, by periodically reviewing the Data Management Plan so that research materials that have been collected can be used effectively by the core research team, collaborators and other potential reusers.

The case study illustrates some of the diversity of research practices in the social sciences, and their influences on the re-usability of video data. To identify relevant support for curation of this research material the report includes a landscape review of tools, resources, and advice services available to the UK Higher Education community and should interest researchers and service providers in this rapidly evolving area. The study also indicates that preservation and curation of video and multimodal research data would benefit if researchers had better-coordinated support for video, across local and national institutional services. Initiatives to publish video corpora are likely to be best led by researchers in the domains concerned, but with coordinated support from institutional and national data repositories in such key areas as storage management, format migration/ transcoding, metadata implementation, ethics and IPR – areas that may already be addressed by institutions' e-learning initiatives. The alternative is likely to see researchers increasingly using commercial web enterprises oriented to 'user generated' video content in ways that neither comply with legal and ethical obligations nor keep data accessible and reusable.

Recommendation 3 – DCC should collaborate with relevant Research Councils, JISC Digital Media and JISC Legal Information to guide institutions, research ethics committees, and researchers on planning and managing the curation of video and multimedia research data.

Recommendation 4 – HEI's should consult researchers on the methodological and technical issues affecting the reusability of video and multimodal data they would want to submit to institutional or subject data repositories, and coordinate the support they provide with the relevant services provided by JISC and other agencies.

1. Video Data in Studies of Interaction

1.1 Scope: Roles and Reusability of Video

The use of video as data for social analysis is increasing although still rare. One factor indicating the significance of video as a research resource is its high profile in e-social science technology development. Another is the availability of relatively cheap consumer oriented tools for digital video capture, editing and distribution. Video is also a costly and complex research resource however - a factor in the relative scarcity of studies featuring video in a significant role. As background to the study, which illustrates some sources of that complexity, this chapter considers why researchers in the domains featured are using and reusing video. These domains are *Video Analysis* in Human Geography and Computing, and *Eye Movement Studies* in Psychology. The chapter gives a broad overview of their recent history, methodologies, typical research topics and settings, aiming to explain why reusing archived video data makes sense in these fields, i.e. its ' curation drivers'.

The focus here is on digital video rather than other forms of moving image[5], and on its use as social science research material. For the purposes of this study, 'social science video data' refers to video that is analysed using some form of social research methodology to analyse human interactions; either 'in' a video recorded setting (e.g. video as observational data), or 'with' a video recorded setting (e.g. video as experimental stimulus). Also the focus is on studies where video plays a significant role alongside other sources, resulting in 'multimodal' corpora. This study reviews *some* research applications of video in Human Geography, Psychology, and Computing, but is not intended to be a comprehensive review of them.

Video is of course used in other ways in the research process as material for analysis and interpretation, for researchers to present their work to their peers or to other audiences, for teaching and learning or public engagement. Some examples are given where relevant to the main focus of the study, which is on video as experimental or observational data.

1.2 Themes and Methods

The SCARP studies have four main themes:

1) *Policy drivers, enablers and barriers:* organisational and institutional factors including different skill levels, preservation policies and arrangements, willingness to use these, and relationships to incentives and reward structures.
2) *Stewardship practices:* how the research process and methods relate to the primary data created and external sources, how these are reused and linked to publications, attitudes to doing this, the usefulness of prior data, and the sustainability of collected digital information.
3) *Tools and infrastructure:* tools and facilities used to collect, deposit, find, cite, discuss and annotate the data, and to ensure persistence and preservation.
4) *Preserving context:* how communities of practice and their knowledge bases can be characterised, and how lineage and provenance is or may be documented.

The methods used in SCARP include:-

- *Semi-structured interviews*; in this case 14 interviews of 45-90 minutes with 10 researchers, and informal meetings with 6 members of Information Services staff with an interest in University of Edinburgh research data services. Interviews with researchers were transcribed and summarised.

---

[5] For example moving images produced in virtual reality environments or in medical imaging fields.

- *Walkthroughs or demonstrations by participants*; in this case a 'guided tour' of a video data archive, by one of the researchers involved in collecting video materials.
- *Observation of meetings and data management activities*; these included non-participant-observation of video data logging and analysis by two researchers, and participant observation in 6 'data sessions' where video clips were presented and analysed. Two of these sessions were recorded on video, to capture interactions between the presenters of video clips and their audience. The aim here was to document typical exchanges between presenters and others less familiar with the clips they were presenting, to better appreciate the role of 'contextual information' in these interactions. Recordings from the data sessions were transcribed and analysed, and were themselves the subject of a data session.
- *Workshop*: study participants took part in a workshop to review curation issues in their research and explore the research data services and roles that could address these.
- *Literature and landscape review* of relevant literature and online resources, including web sites or documentation from the study participants' projects.

1.3 Participants and Stakeholders

SCARP studies have mostly focused on the curation work of a specific group in a given institution or data centre; effectively a sample of one group from a particular discipline. This study instead focused on loosely related social researchers, connected through their use of video data, and methodological preferences rather than through a single project or institutional research group. The rationale for selecting (or inviting) participants was initially no more than that video contributed to their use of a social research methodology, whether in 'pure' or 'applied' social science fields.

Researchers were sampled based on access to personal contacts, or 'snowballing', drawing on their mutual interest in video and/or methodologies for its analysis, with the aim of observing or discussing curation-related activity at different project stages. Participation was also sought through an online survey of academics and researchers across social science departments in the University of Edinburgh College of Humanities and Social Science. This had a small response, bringing two further participants.

The result was a sample of researchers using digital video in a range of projects and research groups. Video data was used more extensively in some groups and projects than others, and had been used for periods ranging from over 10 years to a few months. This meant the participants' views and practices could be contrasted in a variety of interesting ways, including:

- Roles of video data in observational, experimental, or practice-based methodologies
- Research project stage
- Researchers' experience in managing video data

The study involved:

1. Five researchers using ethnographic approaches in observational studies; two in Human Geography, and three in Computing-related projects. All five were loosely affiliated with SEDIT (Scottish Ethnomethodology, Discourse and Interaction) an informal network who meet to collaboratively analyse video data.
2. Three researchers using eye scanning and visualisation methods in experimental studies of visual cognition, a domain in Psychology.
3. Five researchers involved in three projects developing multimodal technologies. One of these participated in a collaborative development project to research the innovation process involved in it, while others were users or developers of tools to support the

analysis of multimodal corpora.

4. Six providers of information services in University of Edinburgh with an interest in video research data (' institutional stakeholders' below) were also involved in the workshop and/or informal meetings.

The SCARP studies have generally involved collaboration with researchers to provide an understanding of existing curation practices, and appreciation of the future options available for developing those further. The present study focused mainly the first two of the above groups; researchers using video analysis in ethnographic and in eye movement studies.

1.4 Data Policy Developments

There are increasingly coordinated efforts by UK Research Councils to ensure that research investigators produce 'data management plans' as a condition of funding (a comparative guide is available from the DCC[6]). Data Management Plans will generally set out how a project intends to deal with the following:

- Legal, rights and ethical issues
- Data collection and development
- Data standards
- Short-term storage and data management
- Access, data sharing and reuse
- Deposit and long-term preservation
- Resourcing

The Economic and Social research Council was among the first to develop a data policy requiring such plans. Bidders for ESRC funding have long been required to review existing data collections and state the contribution that new data will make to the research, and according to the ESRC policy it is mandatory for researchers to offer to deposit qualitative data, such as video, for long-term preservation with the UKDA (UK Data Archive). ESRC data policy is currently under review and likely to be widened in scope.

The EPSRC, which funds computing research on social interaction in areas such as collaborative systems design, identifies as a principle of good scientific practice that "primary data as the basis for publications should be securely stored for an appropriate time in a durable form under the control of the institution of their origin"[7]. This policy may also be developed further to offer researchers guidance and support.

Research on social interaction is also funded by a range of other bodies with no published data policy, such as the Leverhulme Trust. However publicly financed and regulated bodies may well converge around the policy argument articulated by, for example, the OECD[8], that publicly funded research should offer greater transparency and access to research data.

There is recognition of the need to offer researcher incentives for spending effort on outputs such as data archives that are not traditionally assessed and rewarded (Lyon, 2007). However 'sticks' may be easier to specify in policy terms than 'carrots'. The Research Councils UK draft *Code of Conduct and Policy on the Governance of Good Research*

---

6 Curation Policy Guide available at http://www.dcc.ac.uk/resource/curation-policies/ (July 13 2009)

7 EPSRC Guide to Good Practice in Science and Engineering Research (2006) available at: http://www.epsrc.ac.uk/ResearchFunding/GrantHolders/GuideToGoodPracticeInResearch.htm (July 13 2009)

8 OECD Principles and Guidelines for Access to Research Data from Public Funding http://www.oecd.org/dataoecd/9/61/38500813.pdf (July 12, 2009)

Conduct states “responsibility for proper management and preservation of data and primary materials is shared between the researcher and the research organisation”. However it also holds that ‘unacceptable research conduct’ should include “mismanagement or inadequate preservation of data and/or primary materials”, which it defines as including failure to:

- Keep clear and accurate records of the research procedures followed and the results obtained, including interim results
- Hold records securely in paper or electronic form
- Make relevant primary data and research evidence accessible to others for reasonable periods after the completion of the research: data should normally be preserved and accessible for 10 years, but for projects of clinical or major social, environmental or heritage importance, for 20 years or longer.
- Manage data according to the research funder’s data policy, and all relevant legislation.
- Wherever possible, deposit data permanently within a national collection

1.5 Interaction Analysis as Data-Driven Research

According to (Corti, 2007, p.38) “there is not yet a well-developed paradigm, nor an evident blossoming research culture, for secondary qualitative data analysis”. This depends on where the boundaries are drawn around ‘qualitative data analysis’. There are fields that share a focus on interaction, repeatedly analyse examples of it, and have a flourishing tradition of informal data sharing. Corpus-based research and online archives are well established in linguistics and there is some crossover of methods and tools between it and other forms of interaction analysis in social research. This crossover may be fertile ground for developing curation practices.

Much ‘interaction-oriented social research’ is interdisciplinary, makes it difficult subject matter for the analysis of ‘disciplinary’ differences. An overview of the recent history of the fields shows some of their inter-relationships.

1.5.1 Video Analysis in Human Geography and Computing

Ethnographic video analysis takes different forms, informed as it is by the range of perspectives and epistemological assumptions found across the social sciences and humanities. The form practised by researchers in the projects featured in the case study is informed by Ethnomethodology and Conversation Analysis (often abbreviated to ethno/CA), two closely related strands of sociology originating in the 1960’s work of US scholars Harold Garfinkel and Harvey Sacks. Despite having a relatively small number of adherents in social science, ethno/ca has had a significant influence on sociology (e.g. Giddens, 1984) and in other fields concerned with the detail of social interaction. These include educational research as well as geography (e.g. Mondada, 2006) and computing, where ethno/ca has had a particularly strong influence on the fields of HCI (Human Computer Interaction) (e.g. Suchman and Trigg, 2001) and CSCW (Computer Supported Cooperative Work) (e.g. Randall et al 2007). Ethnographic work in these fields has influenced software product design, especially in the practice of requirements analysis and systems evaluation (Jirotka and Luff, 2006), and applications in ubiquitous computing (e.g. Dourish, 2001).

The rationale for using video in studies informed by ethno/ca is also the grounds for its orientation to working from collections of examples, or corpora. Ethnomethodology is the study of ‘ethnomethods’, the methods that we ordinarily use to make sense of the actions of others and make ourselves sensible to them in turn (Suchman et al, 2001) The approach is distinctive for its indifference to the more familiar concerns of sociology with theorising and explaining the details of social interaction in terms of (supposedly) broader theoretical constructs or social forces. It is strongly empirical, concerned with looking at specific strips of naturally-occurring interaction that are recorded and transcribed.

The settings studied for video analysis in geography may be any place where the spatial arrangement of social interaction are accesssible, whether these are in public (e.g. Brown and Laurier, 2005) or workplaces as private as surgical operating theatres (e.g. Mondada, 2007). In computing, settings are usually those where information systems are being used, including control rooms (Luff and Heath, 2002), and healthcare settings (Rouncefield et al, 2003). Unlike other forms of qualitative analysis, the emphasis here is not on extracting quotes and categorising them in order to 'build theory' but on describing how the rich detail of the interaction is 'internally' organised; how its meaning and structure is built on sequentially using the 'shared in common' knowledge of those who are party to it (Randall et al, op.cit.).

It is this orientation to detailed examination of interaction in its situation that leads ethno/ca researchers to repeatedly analyse recordings, and to treat other ethnomethodologically-informed studies as a corpus of examples that may be drawn on. Video has become more readily usable and affordable as a medium for recording interaction, and has lent itself particularly well to studies that, in addition to everyday talk, analyse gesture and bodily orientation towards physical artefacts (e.g. Luff and Heath, op.cit.). Moreover the *reusability* of digital video is of interest:

> *"Video recordings are …persistent and cumulative, they provide researchers with the opportunity of preserving and comparing data across various projects, and allowing a variety of analytic issues and interests to be addressed to the same corpus of materials.."*
>
> *[Video] "…allows researchers to show and share the material on which observations are based, so that they can judge for themselves the persuasiveness and validity of insights and analyses" (Hindmarsh and Heath 2007)*

Data sharing is an important element of ethno/ca studies, through *'data sessions'* that are informally organised through networks of researchers. According to a recent review of data session practice:

> *"There are many ways of holding a data session. In a common format, someone brings in a piece of audio or video data and enough transcripts for the group. The data are played several times, while all observe/listen. Part of a longer interaction may then be singled out for attention and played several more times. The participants take several minutes to write out their observations and analyses of the agreed-upon portion of data. These are then shared with the entire group, and discussion ensues. Data sessions can be open-ended, meaning the participants dive in and see what can be seen in the data. They can be also be didactic or designed to introduce others to a specific phenomenon (in which case the session leader has typically worked on the data beforehand)."* (EMCA, 2007*)*

Transcription has developed as an integral part of conversation analysis. Extracts of transcripts are included in research publications and they, rather than the original recording, are more normally used to warrant claims made about interactional phenomena – although publications increasingly also include sequences of still video frames. Transcription may follow a notation format devised by Jefferson (1984) - included in Appendix 3 – which enables details such as rises and falls in intonation to be recorded, although not all studies (or scholars) need this level of detailed annotation.

In computing, ethnographers may be part of a multi-disciplinary design team, and their outputs will typically include 'co-design' meetings with system designers throughout the project. These will be accompanied by data more conventionally gathered in qualitative research (e.g. field notes, interview records, still and moving images from the setting). Screen capture and keystroke recording has often been used to augment field recordings, and as the range and sophistication of interactive devices has increased, so has the range of 'system logs' to record their usage. This has driven the development of systems such as the Digital Replay System (Crabtree et al, 2006), which builds on previous work to support communication between ethnographers and software designers (e.g. Twidale et al, 1993).

1.5.2 Eye Movement Studies of Visual Cognition

Eye movement research with video differs from ethnographic research in most ways; in the kinds of interaction studied, the data used and the methodology for analysing it. One thing it does have in common with video ethnography is that it is an interdisciplinary field. Eye movements are studied as a method of investigating the cognitive processes involved in attention and scene perception, one of a variety of research topics with a *linguistic* orientation; which includes investigation of both spoken language and reading. Other studies are more *physiological,* including clinical studies of eye movement development and disorders, while modelling of eye movements and the field's history are also topics in their own right (van Gompel et al, 2007).This history is a long one; Wade (1997) traces the first experimental studies to the late 18th century. Notable early researchers included Erasmus Darwin, grandfather of Charles, who developed it as an empirical rather than purely theoretical field of study (ibid.)

Figure 1.1. Lab-based eye tracking apparatus

Advances in eye tracking technology have been key to the development of the field; late 19th century experiments used devices mechanically attached to the eye to record movements, tracing them onto paper and then photographing them. The major advance from this was to photograph light reflected directly from the eye itself, the basis for a revolution in eye movement research in the early 20th century (ibid.)

What these early experiments showed was that contrary to the smoothness of vision that we subjectively experience, eye movement consists of "a series of relatively static glimpses of the world separated by periods of blindness" (ibid. p.52), a phenomenon long exploited by magicians. These glimpses are termed *fixations* and the periods of blindness are rapid jerks of the eye across the field of vision, termed *saccades*.

Photography plays several significant roles in studies of scene perception; both as the basis for eye-tracking apparatus and the experimental stimulus. Photographs have been used as stand-ins for the complexity of real-world everyday scene perception; experiments typically involve research participants looking at photographs. Over the last 30-40 years video has replaced photography as the recording medium, and computing has been developed to track the position of the pupil and the gaze direction using infrared light. Eye-trackers are mostly lab bench –based devices and studies using them require the head to be kept still (as in Figure 1.1). Mobile, head-mounted devices, first used in the 1950's, have in the last two decades allowed more research into the role of head and body movements in eye gaze, and eye movements in action (Wade, 2007).

Tracking eye movements involves sampling the position of the pupil relative to the image being used as the stimulus. Measurements are taken from one or both eyes at a sampling rate of up to 2000 times per second (2000Hz). The pattern of gaze fixations and saccades may then be visualised as in Figure 1.2, which follows the convention of representing saccades as lines joining fixations shown as circles.

Psychologists' interest in these patterns lies in what they may say about human cognition, and particularly the working of perception and selective attention (Henderson, 2007). Much of the work of the Edinburgh group participating in the study, for example, has been directed at understanding the relative importance of 'bottom-up' characteristics of images and 'top-down' cognition, i.e. whether eye movement patterns are best explained in terms of their *visual salience*, the contrast levels, say, or the number of edges present in an image or, on the other hand, in terms of the meaningfulness of the image and objects in it in relation to a task the research participant is attempting to accomplish at the time. Experiments test hypotheses on the relevance of different measures of saliency, using image statistics and task variables gathered by tracking groups of participants as they perform tasks. Accordingly, experiments have so far shown a mixed picture- visual salience matters, but cognitive factors much more so (ibid.).

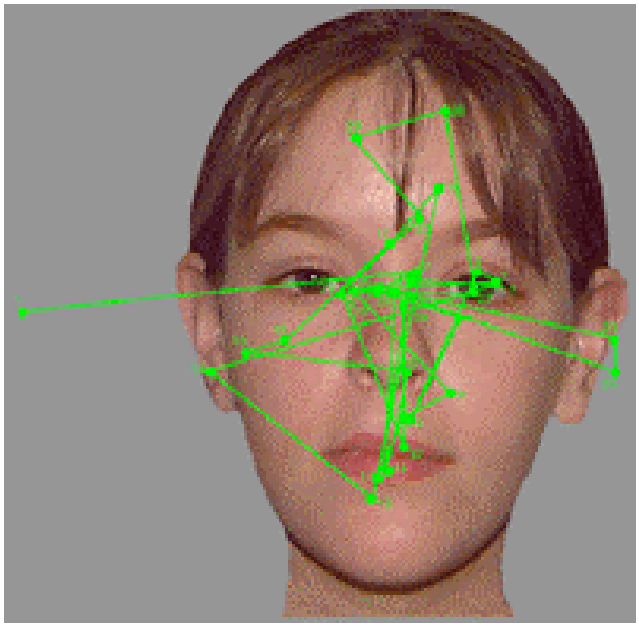


Figure 1.2. Scan patterns in still image

©John M. Henderson and University of Edinburgh, 2009

The use of *video* in place of still images as an experimental stimulus is a recent development, motivated by the greater ecological validity; moving images being closer to our real interaction with the world than still ones:

> *"Ultimately, depictions are stand-ins for the real environment, and what we really want to know is how active scene perception operates in the world itself. Film, video, and virtual-reality environments offer an important middle ground for studying dynamic scene perception while still providing critical experimental control."* (Henderson, 2007, pp.221)

The motivation for a *corpus-based* approach here is that generating a broad range of examples should yield patterns from which hypotheses may be derived, models in this area being under-developed. Corpora have a similar role in studies of eye movements when reading written text. Multimodal corpora combining eye movements, speech, and actions performed in collaborative computer-based tasks have also been used in recent work to explore factors which might benefit the design of interaction with robots (Carletta et al, 2009).

Eye movement research has so far given rise to few published datasets. The *Eye Movement Research Portal* [9]lists two US-based sites publishing work with still or moving images: *DOVES* (a Database Of Visual Eye movementS) is a collection of eye movements from observers viewing still images, which is accessible to researchers on application. The more ambitious (in curation terms) *Online Archive of Scanpath Data* (Scanpaths.org) aims to provide a central online repository of eye-movement data, soliciting these from the research community. The archive provides experiment documentation and aims to develop a "common eye-tracker independent format" usable for developing "new gaze-tracking and visual information analysis algorithms and systems, without having to recreate the experiments carried out to collect the data"[10] The archive currently hosts eleven datasets,

---

[9] Eye Movement Research Portal available at: http://www.eyemovementresearch.com/ (August, 2009)

[10] Scanpaths mission statement, at: http://www.cis.rit.edu/pelz/scanpaths/mission.htm (16 May 2009)

including some derived from experiments using video stimuli and head-mounted trackers.

1.6 Video Recorded Interaction and Context Information

An important and accepted principle in the curation community is that an archived object should be packaged with enough contextual information for it to be reused. The term 'contextual information' is used both generally to refer to all such information or metadata, and more specifically in the Open Archival Information Systems (OAIS) model to refer to:

> *"…information that documents the relationships of the Content Information to its environment. This includes why the Content Information was created and how it relates to other Content Information objects*" (CCSDS, 2002, p. I-8).

The issue here is what this 'relationship between the content and its environment' should be taken to be for interaction research using video recordings. Consider a popular dictionary definition of 'context':

> "*1: the parts of a discourse that surround a word or passage and can throw light on its meaning 2: the interrelated conditions in which something exists or occurs*"[11]

Where should the boundary be drawn between content information (the 'something') and its environment? The first part of the definition suggests that 'context' would include the video content, and the second part that it would consist of links between the video object and researchers' annotations. This leaves open the question of how to distinguish between content, annotation, and contextual metadata in any principled way. That question is not unique to video data of social interaction, but it is an especially pragmatic questions for video curation; since the ability to find and reuse video materials may depend a great deal on how the metadata and annotations associated with it are defined and structured. Also it seems likely that interaction researchers will be interested in both senses of context; in the most granular of details observable or audible and their inter-relationships *within* a recording, and in the inter-relationships *between* the recording, the recorder and the setting.

Interpretive or 'qualitative' social research is typically an exploration of some phenomenon in its context, with more emphasis on this second less granular form of context. Interpretive researchers, especially those who tend towards methodological relativism rather than realism, would see the value in using video as a rich medium for exploring context, but would reject the idea that it *captures* context along with any presumption that video recordings are records of an objective reality (e.g. Pink, 2007). The outputs of qualitative research, particularly in the anthropological tradition of ethnography, are typically written narratives that place a lot of emphasis on the researcher's reflections on how they entered the field setting, experienced it and influenced it. The researcher's presence in the setting is an integral part of what is reported. So for Sarah Pink, one of the foremost proponents of video in ethnography:

> *"…a reflexive approach to analysis should concentrate on how the content of visual images is the result of the specific context of their production and on the diversity of ways that video and photographs are interpreted… they are always representations of the subjective standpoints of the image producer and other viewers, including informants. This has implications for how visual archives are conceived and demands researchers pay attention to the inter-linkages between visual and other (verbal, written) knowledge."* (Pink, 2007, p.114)

Archives should, in Pink's view, enable multiple ways of categorising, sequencing and interlinking images and related texts (ibid.).

This view contrasts with the position that any reuse of qualitative research undermines the

---

[11] Merriam-Webster Online Dictionary: at http://www.merriam-webster.com/dictionary/context (12 April 2009)

principle of reflexivity, that any secondary analysis must be impaired by lack of understanding of the original context, or opportunity to reflect on the researcher's shaping of it. That has been the position of critics of the UK Data Archive's endeavours to promote reuse of qualitative research, who argue that context information is "…different from, unequal to, and a frankly poor substitute for the original interpretive endeavour (however flawed that might be)". (Parry and Mauthner, 2005, p.340)

The long-running methodological debates relating to secondary analysis of qualitative data have been explored in depth, e.g. by Corti (2007). One response is that secondary analysis is a form of *re-contextualisation* rather than an attempt to re-create the primary context; accordingly Bishop (2006) proposes various levels of context information that should be considered a minimum for archiving:

- *Conversation*: transcripts
- *Situation*: characteristics of participants, "e.g. age, gender, marital status, employment status, and so on", date, time, geographic reference, relationships between participants.
- *Project*: description of the originating project e.g. grant proposal, end-of-award reports or working papers, copy of questionnaire if used, consent forms and other materials regarding confidentiality, copy of any interviewer's instructions, details about sampling, data collection methods and data management methods, details of any anonymisation process undertaken, any reports or publications that provide additional information.
- *Institutional/cultural*: depends on the research questions but including relevant media coverage and grey literature for example.

  (ibid.)

While this is helpful in terms of qualitative research generally, it is oriented to research based on interviews with research subjects rather than on records of their naturally occurring interaction; the 'conversation ' level that interaction research is concerned with.

As in qualitative research more generally, the issue of reflexivity is important when considering context at the conversational level. In ethnomethodologically influenced research, reflexivity is seen an aspect of the *recorded subjects'* practice at least as much as an aspect of the *researchers'* practice (see e.g. Macbeth 2001 for detailed discussion). That is to say, researchers are concerned with analysing how those involved in some form of practical action demonstrate subjective (or 'inter-subjective') awareness of self and others in their talk and non-verbal interaction. Reflexivity here means that "actors do not only act but also 'indicate', 'frame' or 'contextualise' how their action is to be understood and how they have interpreted a prior action to which they are responding" (Knoblauch, 2006, p.75). The researcher may be one of the actors, certainly in video recording the unfolding events. Here reflexive action would include the actors' (research subjects') reactions to the camera, or their interactions with the researcher or camera operator. Researchers may therefore need to account for decisions on whether to include these in the corpus of material resulting from the study (Laurier, 2008).

An ethnomethodological perspective on context information given by Dourish (2004) is helpful to understanding the needs of this domain. Dourish questions the assumption in the field of ubiquitous computing that context is a delineable, stable form of information that is separable from the activity or content it is deemed to relate to. Instead he proposes that:

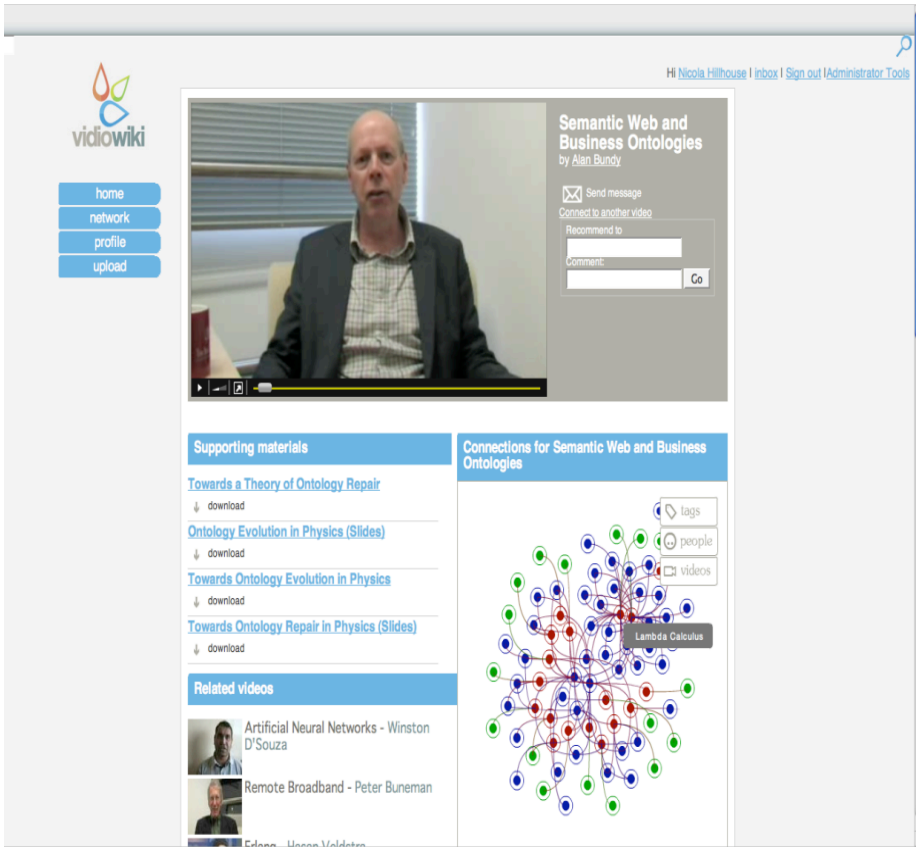
> *"…contextuality is a relational property that holds between objects or activities. It is not simply the case that something is or is not context; rather it may or may not be contextually relevant to some activity… the scope of contextual features is defined dynamically… context is particular to each occasion of activity or action…context isn't just 'there', but is actively produced, maintained and enacted in the course of the activity at hand*" (Dourish, 2004, pp.22)

Taking this perspective, an archival system appropriate for video interaction research would support a dynamically evolving relation between context information and archived video

object. If the primary researcher records at first some minimal 'situational' details of the recording, and in replaying it interprets and documents how the recorded interaction unfolds, these annotations would also be treated as context information, *and any others made subsequently*. This view is reflected in some recent archival models for curating video data in which context information is built up incrementally as archive users interact with the archived object by selecting new sequences to annotate. These models and others are potential drivers of video data reusability, considered below.

Research communication with online video

A growing number of cross-disciplinary web platforms are being used by social researchers to communicate about their research beyond the reach and form of the traditional print journal. Some such as *Videolectures.net* take the form of online lectures, aligning video and presentation slides with academic text, and are subject to some form of peer review. Others such as *ITunesU* are educational channels of more general-purpose multimedia publication platforms.



A recent addition is *Vidiowiki.com*, intended to promote inter-disciplinary contact between researchers by identifying similarities in the content of their video presentations. Users upload video presentations and related textual summaries. Any effort they put into maintaining a map of their topics of interest is rewarded with suggestions of connected users and topics.

1.7 Drivers of Video Data Reusability

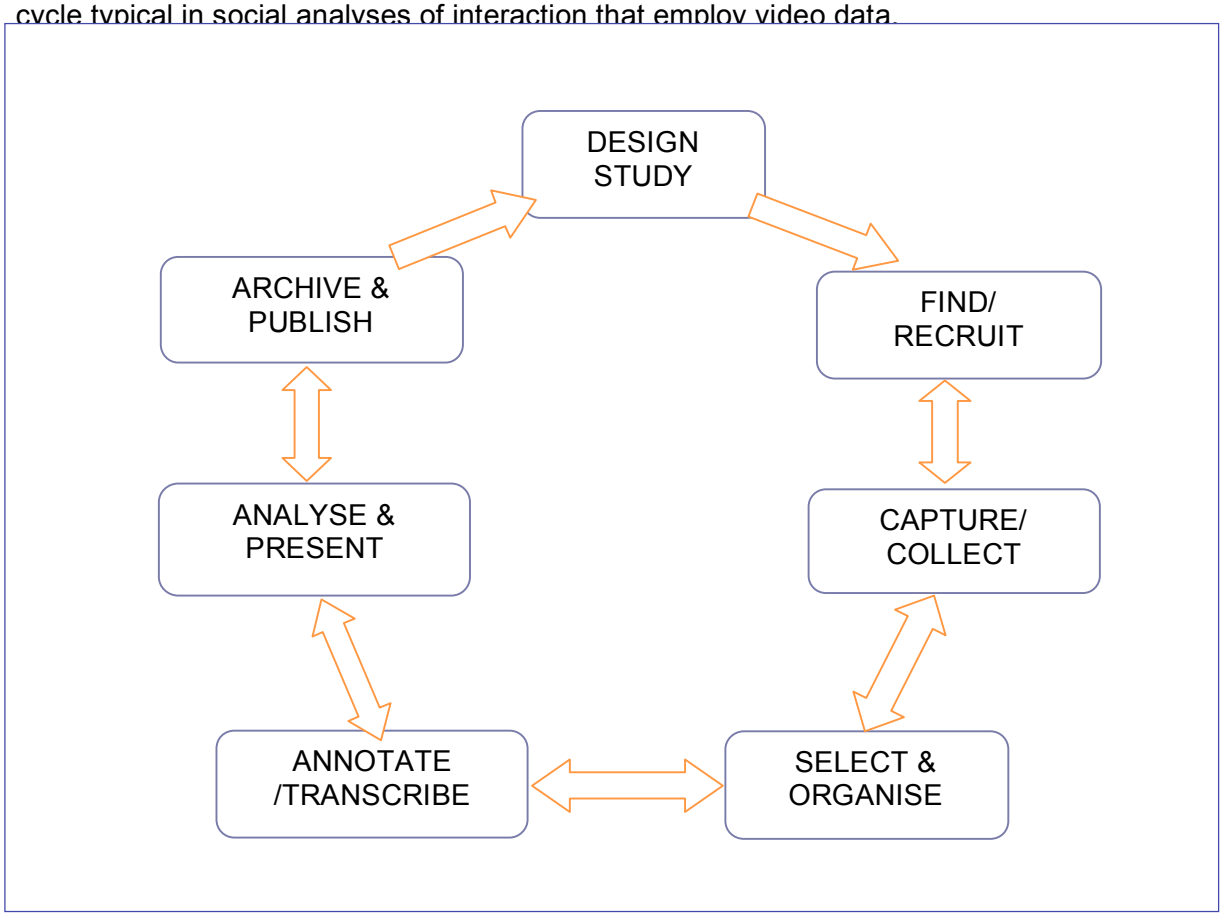
Creating, sharing and commenting on online video are now commonplace social activities. While web video platforms are being adopted for research communication (see box above) social research uses of video are relatively uncommon. They are nevertheless also driven by

increased availability of video data and the tools to work with it. This section considers these as 'drivers' alongside recent development in e-social science.

Enabling the exchange and reuse of social science data has been one concern of the UK e-Science programme, and uses of video as a research resource have featured prominently in that programme. Various e-Social Science projects in the UK and internationally have piloted infrastructure to support the use of video as research data and as a communication tool. Fielding's 2003 study for the ESRC, which informed the e-Social Science programme, claimed that:

> "*Digital video streaming enables real-time interaction analysis, not only in workplace studies, but in linguistic research concerned with language use in context, work on human/computer interaction (including system design), observational classroom/training studies, and visitor studies (e.g. responses to museum exhibits)*". (Fielding 2003, p.4).

Fielding's report presaged a range of projects to develop e-social science tools and infrastructure, with interaction analysis featuring prominently as a multi-disciplinary approach, and real-time video delivery as a means to support it. To briefly review this and other 'drivers, enablers and barriers' to reusability, Figure 1.3 below gives an overview of the research cycle typical in social analyses of interaction that employ video data.



The research cycle in Figure 1.3 shows seven stages, although as the direction of the arrows indicates there is much overlap between these stages. The main factors driving interest in curating video and multimodal data coalesce around three parts of the cycle:

- *Finding, Capturing and Collecting* new capture technologies, new sources of digital video and multimodal data, and new means to organise this data present challenges in identifying what to preserve.
- *Selecting, Annotating and Analysing*: an increasing availability across several research communities of tools to annotate multimodal data and make these annotations web-accessible.

- *New approaches to archiving and disseminating interaction research*: together with the changes in tools and technologies for working with video, new archival models for multimodal data analysis and new channels for research dissemination are enabling inter-disciplinary collaboration around video and other recordings of events on a shared timeline.

1.7.1 New Resources for Capture and Collection

Several factors driving research uses of digital video are obvious; camcorders and video editing software marketed to the consumer (or 'prosumer') have become affordable within small research project budgets, and typical desktop computers used by researchers are capable of running these. Meanwhile web-based digital video has within the space of a few years become a pervasive feature of ordinary social life, joining a wide range of online activity that is potential material for research on social interaction.

The extent of material accessible to researchers is vast; it has been reported that the video footage uploaded to the YouTube web platform in the first half of 2008 exceeded the combined broadcast output of the four major US TV channels over the previous 60 years (Wersch, 2008). On a slightly more modest scale, web video material specifically intended for UK Higher Education has emanated from the JISC digitisation programme, which since 2006 has funded projects to make visual resources web-accessible for educational use. Major JISC funded sources include the *Film and Sound Online* and *News Film Online* services, provided through EDINA[12].

For several years the pace of change in capture technologies has also been very rapid. Earlier sections have already pointed to the significance of new eye-tracking technologies for eye movement research, and of new forms of digital interaction as a source of data in social computing research. While digital video itself has become easier to acquire, there has been a proliferation of technologies for recording it at the point of capture, for capturing it at higher resolutions, for storing, editing and organising it.

Changes in *recording* technologies include the availability of 'high-definition' (HD) video camcorders for the consumer market, with wider take-up of compatible TV sets and broadcasting. HD video provides roughly double the image resolution of the PAL standard long established for European broadcast TV and video, and a wider 'letterbox' aspect ratio (16:9 as opposed to 4:3). For immediate *storage*, mini-DV (digital video) tapes have been established since the mid 1990's but camcorder users now have the option of equipment with hard disk or flash memory storage.

Uncompressed video at standard definition consumes around 93Gb per hour[13]. However camera technologies generally compress at the point of capture; the DV format compresses 'raw' video by a factor of 5. While HDV provides higher resolution than DV it also reduces the captured information by a further factor of 4 using MPEG-2 compression. HD camcorders also use various other proprietary formats. Interaction researchers therefore have to balance the advantages of HD for capturing gestures or facial expressions with uncertainty in future software support for rapidly changing formats (Trilsbeek, 2007).

The recent JISC study on *Significant Properties of Moving Images* (Coyne and Stapleton, 2008) addresses the issue this rapid change poses for curation; it brings obsolescence and it follows that to preserve the original object means identifying which of its characteristics need to be preserved through further changes. The report takes the OAIS model, with its concept of the *Archival Information Package*, as its starting point. This OAIS schematic representation of an AIP is as shown in Figure 1.4.

---

12 EDINA available at: http://www.edina.ac.uk

13 Example from JISC Digital Media at: http://www.jiscdigitalmedia.ac.uk/movingimages/advice/choosing-a-digital-video-file-type/ (April 20, 2009)

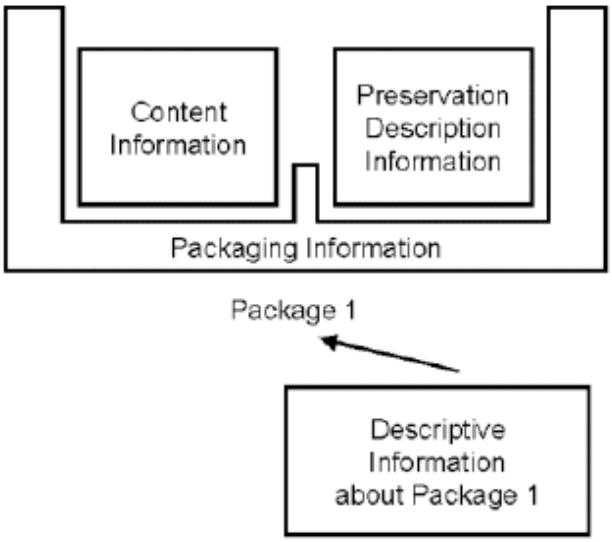


Figure 1.4 Archival Information Package (OAIS CCSDS, 2002, Fig. 2-3)

Archival Information Packages (AIP) are derived from *Submission Information Packages* (SIPs) provided by depositors, and translated into *Dissemination Information Packages* (DIPs) that meet the needs of archive users, according to the OAIS functional model shown in Figure 1.5.

An AIP for video, which in this report is taken to include the moving image together with any audio stream, would be expected to include the data object encoding the content as a sequence of bits, plus *representation information* or technical metadata that is required to make the bitstream retrievable as a meaningful digital object, and *Preservation Description Information* needed to preserve that object. Preservation Description Information comprises *provenance* (the history of actions performed on the data, by whom and when); *reference* (information resulting from applying a unique identifier from an identification scheme), *fixity* (information documenting the authenticity mechanisms for the data and *context* information (documenting the relation between the content and its environment, including why the content was created).

In practice metadata standards commonly apply different terminology, which nevertheless should overlap with most of the above to support[14]:

- *Descriptive Metadata* enabling identification, location and retrieval of information resources by users, often including the use of controlled vocabularies for classification and indexing and links to related resources.
- *Technical Metadata* describes the technical processes used to produce, or required to use a digital object.
- *Administrative Metadata* is used to manage administrative aspects of the digital object such as intellectual property rights and acquisition. Administrative Metadata also documents information concerning the creation, alteration and version control of the metadata itself. This is sometimes known as meta-metadata!
- *Use Metadata* defining user access, user tracking and multi-versioning information.
- *Preservation Metadata*, documents actions which have been undertaken to preserve a digital resource such as migrations and checksum calculations.

---

[14] From Description of a Metadata Standards DCC Standards Watch February 2007 available at: http://www.dcc.ac.uk/resource/standards-watch/what-are-metadata-standards/ (April 13, 2009)

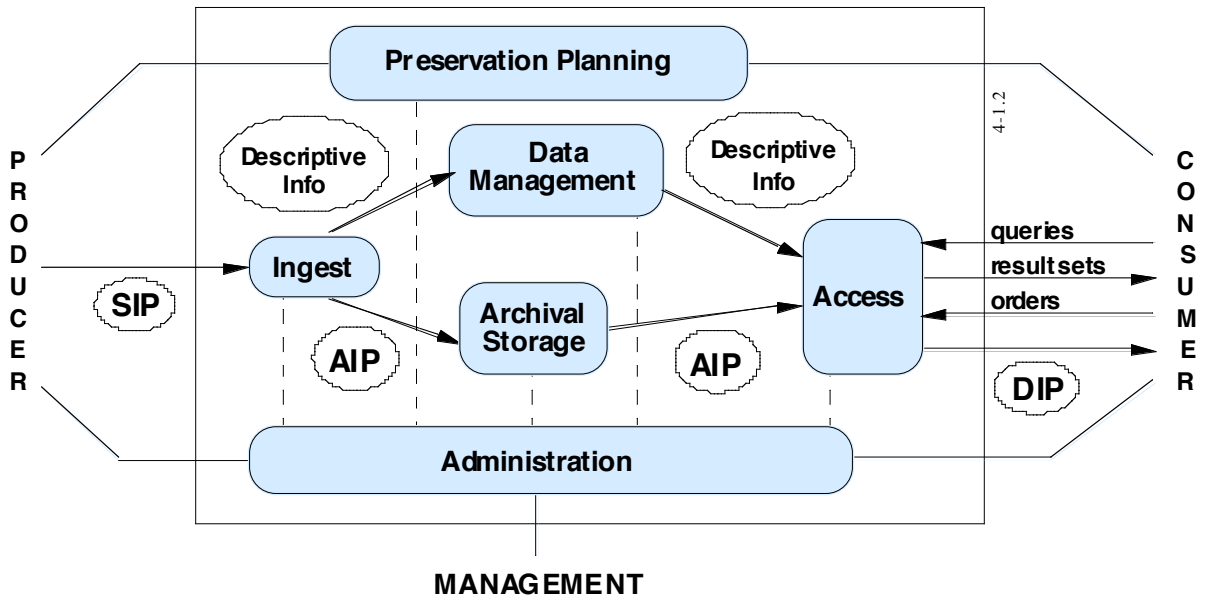


Figure 1.5 OAIS functional model (CCSDS, 2002)

The term '*significant properties*' refers to information needed to perform interaction between the data object and the technology needed to render it (Coyne and Stapleton, ibid.) Significant properties includes technical metadata and is a subset of the 'representation information' in the OAIS schema. For video, significant properties comprise:

Content: Number of image and audio streams, their length

Context: Title, creator, date of creation, provenance

Rendering: Frame width and height, bit depth, colour, aspect ratio, frame rate, compression type and ratio, codec and scan type

Structure: Relationship between audio and image streams; relationship between metadata and bit streams

Preservation begins with the capture or collection of video material since some of the significant properties can be retrieved from the wrapper files (e.g. .mov, .avi) created by capture and editing software. The quality of the video and audio content information retained, and the ability to retrieve and reuse the automatically created metadata depend on using appropriate standards. These are considered later, and discussed in depth in *Significant Properties of Moving Images* (Coyne and Stapleton, ibid.) and the *Moving Images and Sound Archiving Study* for the Arts and Humanities Data Service (Wilson et al 2006).

1.7.2 Interdisciplinary Tools for Annotation and Analysis

Software tools to inter-relate media files and link these to research notes have rapidly emerged from the qualitative analysis and linguistics communities. Some tools originating in specific fields of research or professional practice have had wider take-up, while others have been consciously designed to bridge disciplines. Other more general-purpose technologies potentially support research; a growing number of web 2.0 video annotation tools, together with methods for extracting text from video, enable effective sharing, browsing and searching

of video and multimodal resources.

Considering general purpose tools first, various web-based video *search engines* are now available, employing automated feature extraction and classification techniques based on analysis of video streams, or using speech recognition techniques on audio streams. Such systems are typically 'trained' from a manually annotated set of examples, or use user-defined tagging to improve the effectiveness of automated techniques. Examples include *Blinkx*[15], *Truveo*[16] and IBM's *Multimedia Analysis and Retrieval System* (IMARS)[17]. A review for the AHRC by Marsden at el (2007) outlines further examples of these and their application in humanities research.

Also relevant here is the recent upsurge in web 2.0 tools offering free-form collaborative annotation of video clips. These include for example the recently launched annotation service in *Youtube*[18]. *Vimeo*, a platform for sharing high-resolution clips between members of self-maintained groups or 'communities' also allows members to comment on these videos[19]. The *Voicethread* platform[20] offers users the capability to add multimedia comments on still images or multimedia clips, while *Kaltura* provides wiki platform extensions that enable collaborative online editing of video clips embedded in wiki's[21]

Annotation has recently emerged as a focus of e-Science and curation research (e.g. Buneman, 2005) and support for it is a key element of the infrastructure for research uses of video. Less than a decade ago a UK-wide academic infrastructure for video data management was proposed, operating on client-server lines to provide video storage and file format migration (transcoding) services (Shotton et al, 2003). While storage and network bandwidth have become cheaper and more available, online video players now allow individuals to transcode clips between distribution formats in real-time. In keeping with this, institutional infrastructure for online video has re-emerged as an issue, with a focus on promoting course content and on research communication, as in the JISC *Steeple* project[22].

Collaborative manual annotation and retrieval of *research materials* remains a challenge however. One aspect is the integration of ontologies with video annotation tools, and the *PolicyGrid* architecture (Edwards, 2009) is seen as having potential application to time-aligned resources (e.g. synchronised video, audio, and transcripts) (Shotton et al, 2008).

The capability to annotate specific points in a sequence of video or other timed data, as required for interaction research, has not so far been provided in general purpose web 2.0 video sharing platforms. MacWhinney et al (2005) envisage the *W3C Annotea* web-based annotation protocol being used for interaction research purposes, and more recently the *Vannotea* project at the University of Queensland has used this protocol in their system which provides real-time collaborative annotation, search and retrieval of high-quality multimedia streams (Shroeter et al, 2006).

Several web-based annotation tools intended for e-learning purposes have potential application to interaction research. MacWhinney et al (ibid.) discuss *Project Pad*, a tool for web-based multimedia annotation developed at Northwestern University[23]. The *Diver* tool

---

15 Blinkx: http://www.blinkx.com/ (March 16,2009)

16 Truveo: http://www.truveo.com (March 16, 2009)

17 IMARS: http://www.alphaworks.ibm.com/tech/imars (March 16, 2009)

18 Youtube annotations: http://www.youtube.com/t/annotations_about (March 16,2009)

19 Vimeo: group-based video sharing available at http://www.vimeo.com (March 16, 2009)

20 Voicethread: multimedia annotation available at http://voicethread.com (March 16, 2009)

21 Kaltura collaborative online editing: http://www.kaltura.com/devwiki/index.php/Main_Page (March 16, 2009)

22 Steeple project: http://steeple.oucs.ox.ac.uk/index.html (July 31, 2009)

23 Project Pad: http://projectpad.northwestern.edu/ (March 16, 2009)

developed at Stanford University's Center for Innovation in Learning[24] provides annotation or logging in a novel interface, Designed for reflective analysis of learning, e.g. in classroom settings, the tool creates a panoramic overview of an imported video sequence. *Diver* enables the user to create an annotated path or 'dive' through the sequence, share 'dives' on a *Webdiver* site that in turn allows others to comment. The JISC-funded *Synote* project at Southampton provides a web based tool for creating 'synchronised bookmarks' that can contain notes and tags synchronised with audio or video recordings, transcripts and slides/images, and can be used to find and replay parts of the recordings[25].

Tools for video annotation in interaction research have until recently been standalone tools, designed to fit the categories, methods and methodological assumptions of their originating field. Some of these tools have been taken up more broadly in interaction research communities. The widely used video transcription tool *Transana* was originally developed to meet the needs of educational researchers but is widely used in other fields for qualitative analysis of video material. The software supports the Jeffersonian transcription convention used in conversation analysis. It has also extended the notion of 'transcript' to include other time-aligned texts such as translations and field notes or 'logging'. *Noldus Observer* a video annotation tool developed for behavioural studies is also now more generally used. Various other tools are identified by Brundell et al (2008).

Tools developed specifically to annotate multimodal corpora for linguistics purposes have also been taken up in related sociological areas of interaction research. For example *ELAN*, developed by the Max Plank Institute for Psycholinguistics, has been used to encode ethnographic data. The *NITE XML Toolkit* (NXT) developed for annotation in computational linguistics is also being used in the AMI corpus to meet the analysis needs of organizational and social psychologists (Carletta, 2006).

Linguistic annotation structures are complex, so cross-disciplinary annotation of a multimodal corpus requires annotation tools that are flexible enough to allow different media streams (also known as 'signals' in multimodal research) each to be labelled or commented on using tags from different coding schema. For example NXT provides the capability to define coding schema for annotations, inter-relate these to each other, and to the start and end-times of synchronised media streams. NXT allows sophisticated Boolean searching and filtering of annotations by their type and timing.

Computer-aided qualitative data analysis (CAQDAS) tools such as *Nvivo* [26] have traditionally operated on a stand-alone basis, tending to have limited support for multimedia (Friese, 2004). CAQDAS tools have been developed to support annotation of segments of texts, such as interview transcripts. Their capabilities are oriented to qualitative analysis that involves disassembling research materials into constituent parts, sorting and coding them, and re-assembling them into a meaningful narrative. Recent versions of *Nvivo* and *Atlas.ti* [27] have introduced the ability to annotate video at a granular level. CAQDAS tools however lag behind multimodal corpus annotation tools in the capability to synchronise data sources on a time-line (Brundell et al, 2008).

Among the barriers to exchanging CAQDAS coding schemes – and therefore to reusing them within or across disciplines, is that CAQDAS software rarely exports data in XML. A pilot project run by the UK Data Archive - *DEXT* – sought to overcone this by producing an XML schema called QuDEx to describe qualitative data annotations and their structure. This followed work in the SQUAD project that used NXT to partially automate some annotation

---

24 Diver project: http://diver.stanford.edu/what.html (May 23, 2009)

25 Synote project: http://www.synote.org/synote/ (May 23, 2009)

26 Nvivo: http://www.qsrinternational.com/products_nvivo.aspx (May 23, 2009)

27 Atlast.ti: http://www.atlasti.com/ (May 23, 2009)

tasks, such as transcript anonymisation[28]

In the UK e-Science community, a series of projects – *Vidgrid, MiMeG* (Mixed Media Grid), and *DRS* (Digital Replay System) have developed prototype tools to support the "widespread informal collaborations between video-based researchers, collaborations that involve co-present and collaborative data analysis workshops", given that "some prominent approaches to video-based research prioritise data sessions as a way of working, emphasising the value of real-time discussion and analysis of data. (Fraser et al, 2006).

The Vidgrid and MiMeG projects adopted a view of annotation as a feature of *real-time* data analysis, an integral part of discussing and gesturing "over and around video data and associated materials" *rather than* "as an activity to be performed for purposes of managing data and metadata" (ibid.). Metadata were limited to the log-in details used, the project and associated file names, and the main features of MiMeG were capabiliies for remotely located data session participants to work together. Co-located participants shared a projected screen and used a mouse, headphones and microphone, plus an electronic pen to control and annotate the projected video. Remotely, the participants could each see the others' annotations. However trials founds that participants found it difficult to coordinate their responses as they could not see each other in the action of making these annotations (Fraser et al, ibid.)

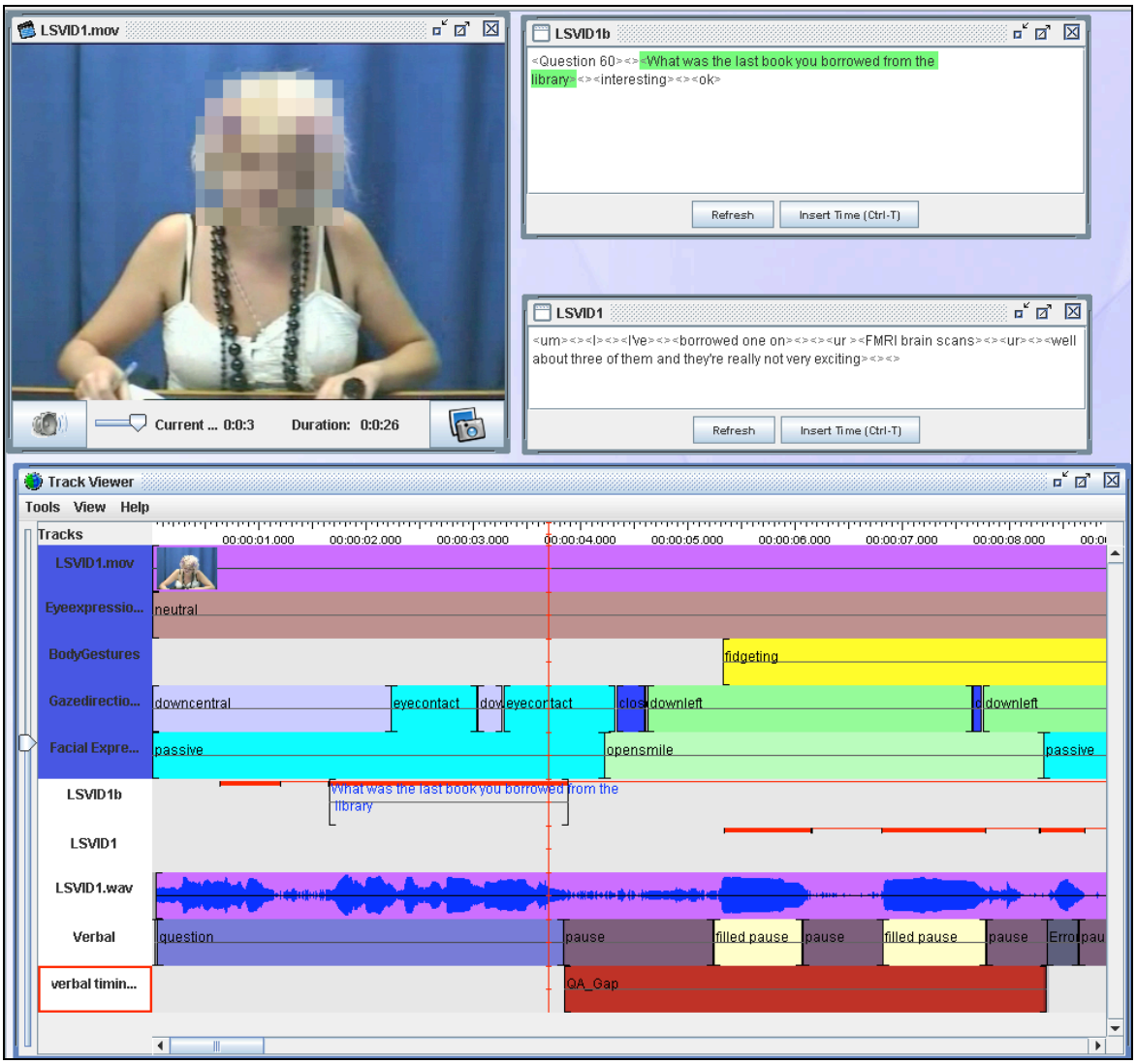


Figure 1.6 Digital Replay System (Brundell et al, 2008)

28 UK Data Archive DEXT and SQUAD projects at: http://www.esds.ac.uk/qualidata/online/about/dtd.asp

The *Digital Replay System* (DRS) has adopted a different tack, focusing on annotation of video alongside other time-based data such as system logs, to analyse how different modes of interaction are temporarily related. DRS replays synchronized video and associated data that is 'internal' to the interaction, together with transcripts, annotations, and any coding schemes used for this (see Figure 1.6). The system also provides capabilities to manage the corpus of material associated with a project, allowing for multiple "analyses" or subsets of these (Brundell et al, 2008:2) i.e. an 'analysis' is a corpus with shared provenance.

The attraction of video data here is to enrich the analysis potential of recorded human interaction, which until recently has been largely textual:

> *"…while the analysis of large-scale text corpora can provide insights into language patterning and can help establish linguistic profiles of particular social contexts, it is limited to the textual dimension of communication. Communication processes are multi-modal in nature and there is now a distinct need for the development of corpora that enable the user to carry out analyses of both the speech and gestures of the participants in a conversation, and of how the verbal and non-verbal complement one another."* (Carter and Adolphs, 2008, p.275).

Like the Max Plank Institute's ELAN tool and the AMI Project's NXT tool, DRS provides for multiple annotation "tracks" ('tiers' in ELAN) to be applied to a timeline, so that researchers can apply different analytic schemes to various data sources gathered over the same recorded time period; for example linguists, psychologists and e-learning specialists collaborated in one of the 'driver projects' that DRS was developed to support (ibid.). Such capabilities for inter-disciplinary analyses are being used in repositories consciously designed for inter-disciplinary use, which we return to below.

1.7.3 New Archival and Communication Models

Web video may have spurred multimodal archival resources aimed at *inter-disciplinary* work in areas of psychology, linguistic and learning research. The idea of video as a medium for inter-disciplinary work has a history though. For example the *Diver* platform has been envisioned as the basis for 'digital video collaboratories' influenced by the 'Interaction Analysis Lab' work of Jordan and Henderson in the early 1990's (Pea, 2006).

The potential for *longitudinal* qualitative studies to reuse multimedia data is currently being explored in, for example, the *Timescapes* project, the first large-scale qualitative longitudinal study in the UK (Holland et al, 2006). The project uses video interviews in tracking individuals and family groups over time to document changes and continuities in their relationships and identities. A data repository, the Timescapes Archive, has been set up as a specialist satellite of UKDA Qualidata[29]

Corpus-based approaches to interaction analysis represent a different form of reuse and digital technologies have a well-established role in supporting these. Archiving of video in multimodal corpora is relatively well established in linguistic domains of Humanities research. There is already some overlap in the use of analytic tool use between, for example, the domains of socio-linguistics (Humanities) and conversational analysis and ethnomethodology (Sociology).

Web-based frameworks for browsing annotated corpora are a more recent development. These do more than enable the online searching of metadata; providing web access to transcribed and annotated recordings. The longest established archive is *Talkbank*[30] at Carnegie Mellon University, which aims to advance the development of standards and tools for creating, sharing, searching, and commenting upon multimodal records of human and

---

29 Timescapes Archive available at: http://www.timescapes.leeds.ac.uk/the-archive/ (23 April 2009)

30 Talkbank: http://talkbank.org/ (23 April 2009)

animal communication. Talkbank's CHILDES child language development component provides a browsable database of annotated transcripts, with synchronised audio and video clips. Contextual metadata is fairly minimal, comprises language, participants, location, situation and activities, each typically described in one or two lines of text.

Multimodal corpora are increasingly the focus of inter-disciplinary projects aiming to enable their re- analysis from a variety of perspectives by providing web access to data. The *Social Informatics Data Grid* (SIDGRID) at University of Chicago described as:

> *"…at the intersection of computer vision, psycholinguistics, cognitive neuroscience, neuroscience, psychology, linguistics, education, anthropology, speech and language processing, and high speed computing and networking", aiming to enable researchers to "capture multimodal behavior in real-time at multiple levels simultaneously, and then to store and analyze different data types (e.g. voice, video, images, text, numerical) in a distributed multimedia data warehouse that employs web and grid services to support data storage, access, exploration, annotation, integration, analysis, and mining of individual and combined data sets.*" (Bertenthal et al, 2007)

The *AMI Meetings Corpus* offers linguists and organisational psychologists streamed video recordings of meetings, allowing users to view selected combinations of camera views. Transcripts, annotations and metadata can be downloaded for use with annotation browsers on the user's computer. Another example is CLAPI a web-accessible archive of (French language) naturally occurring social interaction data at the University of Leon, comprising around 120 hours of audio and video (Bert et al, 2008).

The key capability afforded by web-based multimodal corpora such as AMI and CLAPI is to visualise *multi-disciplinary* views of a time-line in an audio-visual sequence, where these views correspond to multiple 'tiers' or 'tracks' created using an annotation tool such as ELAN, NXT, or the DRS product of the UK NCeSS programme. These however still require the researcher to download the corpus or item of interest, in order to see how it has been annotated.

The use case for *web* browsing is to allow users to quickly inspect annotations together with their target sequences of interaction, before choosing whether or not to download for more accurate analysis (Kemps-Snijders et al 2008). Archives for interaction research do not as yet provide web-based capabilities for *secondary* annotations to become part of the archived record, although that is envisaged as part of the Max Planck Institute's *Annex* framework (Berck and Russel, 2006).

Another more general model supporting archiving of secondary annotations is the Contextual Information Framework (Figure 1.7) proposed in the *VidArch* project (Lee, 2007). This supports the principle that contextual information should continually evolve, with objects being linked to successive versions with the user community's annotations and usage data.

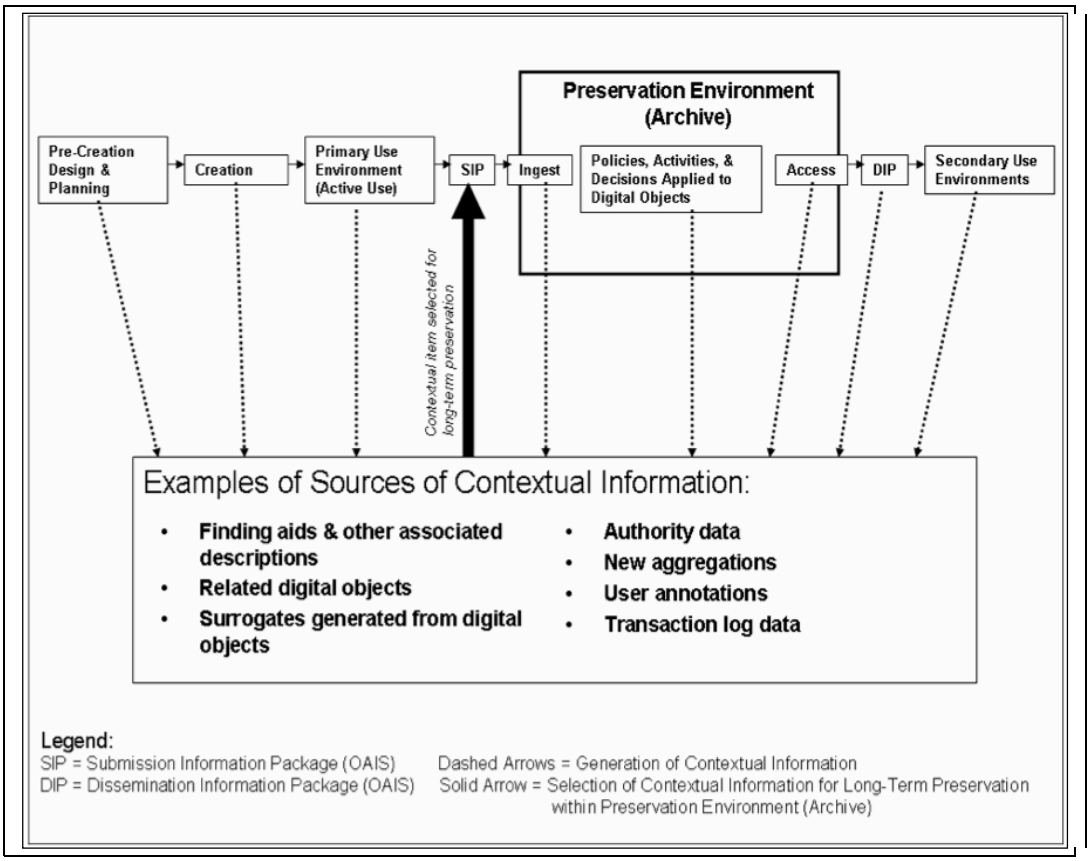


Figure 1.7 Contextual information in the life of a digital object (Lee, 2007)

2 Challenges in Curating Video Corpora

The researchers featured in this study adopt 'data driven' approaches to working with video that the previous chapter linked to their disciplinary traditions and inter-disciplinary outlook. Curation of that data is challenging however for methodological and technical reasons this chapter describes. Following an overview of the research groups and projects that featured in the study there are two main sections. Section 2.2 describes the diversity of roles that video plays in the research groups involved in the study. Section 2.3 deals with the uncertainties that make planning for curation difficult; including the boundaries of the contextual information that would support data reuse in interaction research.

2.2 Participation in the Study

2.2.1 Ethnography in Human Geography and Computing

SEDIT (Scottish Ethnomethodology, Discourse and Interaction) is an informal network of researchers from a variety of social science and humanities disciplines, including International Relations, Psychology, Linguistics, Health Studies, English Literature and Sociology, as well as Human Geography and Computing Science. SEDIT Meetings are organised as 'data sessions', where researchers collaboratively analyse video or audio clips with transcripts on topics reflecting the inter-disciplinary mix. The group meets weekly in the University of Edinburgh's School of Geosciences, regularly involving people from Glasgow, Strathclyde and St Andrews Universities and with occasional visitors from English, European and US institutions.

Two people from this group were involved through repeat interviews and observation; a Senior Research Fellow and Research Assistant in Human Geography, part of the Institute of Geography at Edinburgh. Three researchers from Computing research groups were also interviewed. These were two Research Fellows from Edinburgh and Glasgow Universities and a Research Assistant from St Andrews University.

Human Geography

Video-related projects discussed here included:

Habitable Cars: This 30-month ESRC-funded project[31] was completed in 2007 and examined the "contemporary spatial organisation and practical arrangement of collective car travel and car sharing". Digital video was the main data source used, and among the project's aims were to "secure the use of video as robust evidence in social and cultural research".

Assembling the Line: a three-year ESRC-funded project (2007-2010)[32] applying video ethnography to investigate amateur and professional video editing. More specifically the research examines the nature of editing as work, the skills involved in analysing and assembling materials, and how competence in editing is acquired in amateur and professional settings. It also considers how editors' video analysis practices compare with social scientists'.

Concierges at Work: Maintenance and the Making of Residential space: using an audio-visual data corpus documenting the work of concierges in a high-rise estate, compiled for an

---

[31] Habitable Cars: http://www.esrcsocietytoday.ac.uk/esrcinfocentre/viewawardpage.aspx?awardnumber=RES-000-23-0758 (accessed 21 April 2009)

[32] Assembling the Line: http://www.esrcsocietytoday.ac.uk/ESRCInfoCentre/ViewAwardPage.aspx?ts=2&data=z8HSvl3fWwVY2sDo4JNP8iOLJQdQnq85XXHZBtq2B42FRD%2FIlNVd6r%2BWHfqR%2FxHWTxYOt1BSx7MWmIxu%2FM1BIG%2B3N%2BHD89pW%2F1O02wttQxo%3D (accessed 21 April 2009)

AHRC funded research project '*Difference and Repetition: the residential high-rise as a global form*' [33] completed in 2007. The research explores how everyday, on-the-ground tenant and building support is delivered through different styles of concierge servicing, and links this to theoretical questions about the need for social scientific attention to the routine work of maintenance and repair.

Computing

Computing researchers participating in SEDIT were from various Scottish institutions. Each was employing an ethnomethodological approach, using video analysis to help examine social interactions that technologies were being designed for or used in. Projects concerned a wide range of topics including computer programming itself, following sports events, and radiological training e.g.:

Distributed Intelligent Learning Environment for Mammographic Screening: a three year EPSRC-funded project (2007-2010) [34] is developing and testing an 'Intelligent Tutoring and e-Learning Environment' to investigate how this may be used effectively in training radiologists who are intending to specialize in mammography. An important element is observational work to understand the training practices and their context. Coupled with the close participation of trainees and mentors in the development of the e-Learning environment, the observational analysis using video and other records is intended to ensure this matches their needs, and is easy to understand and use.

2.2.2 Eye Movement Research

The Visual Cognition Group in the University of Edinburgh's Psychology department is building a corpus of videos clips and eyetracking data from people watching the clips, in the two-year Leverhume-funded *Dynamic Images and Eye Movements* (DIEM) [35]. This is investigating how people see and understand the visual world as depicted in film and video. An exploratory approach is being taken, using a wide range of video genres as experimental stimuli, including recorded meetings, instructional and demonstration videos, TV coverage of debates, and CCTV images. Each resulting corpus will feature different aspects of attention to dynamic movements including faces during conversation, social gaze cues and gestures. These will be used to identify eye-movement patterns and a selection then used in more detailed experiments and image analysis.

2.2.3 Innovation Studies

Video was being used in a study of technology design and usage, being conducted in the Institute for the Study of Science, Technology and Innovation (ISSTI), University of Edinburgh. Participants in the *Branded Meeting Places* project, funded jointly by AHRC and EPSRC, have used recordings of formal and informal meetings in a range of environments to explore and develop strategies for improving the technologies that support them. The methodology involved "reflective engagement between designers, users and meeting participants" and research evidence included "several hours of video blogs summarising a series of project meetings"[36]

2.2.4 Tools and Resources for Analysing Multimodal Interaction

Later in the project, researchers developing tools and resources for use in multimodal

---

[33] Concierges at Work: http://www.geos.ed.ac.uk/homes/istrebel/projects-ongoing.html and its parent the Highrise project http://www.ace.ed.ac.uk/highrise/ (accessed 21 April 2009)

[34] Distributed Intelligent Learning Environment for Mammographic Screening: http://gow.epsrc.ac.uk/ViewGrant.aspx?GrantRef=EP/E033490/1 (accessed 21 April 2009)

[35] The DIEM Project: http://www.psy.ed.ac.uk/diem/ (accessed 21 April 2009)

[36] Final Report Branded Meeting Places: Ubiquitous technologies and the design of places for meaningful human encounters: available at http://ace.caad.ed.ac.uk/Branded/ (accessed 21 April 2009)

interaction research were interviewed. They were members of two projects:

The *Augmented Multiparty Interaction* (AMI) Project is a European-funded multi-disciplinary consortium, including the Human Communication Research Centre (HCRC) in the University of Edinburgh. Aiming to develop new forms of 'meeting browser' that make available to group members a historical record of their meetings, AMI has produced the *AMI Meeting Corpus*, a publicly available set of multi-modal recordings comprising multiple audio and video sources, plus outputs from projectors, whiteboards and other devices.

The *Digital Records for e-Social Science* or DReSS project at Nottingham University Department of Computer Science has developed the DRS (Digital Replay System) a tool for synchronization, replay, and analysis of audio and video recordings. The DRS tool allows these to be combined with records of interaction within computational environments, such as SMS messages, GPS data and data from body sensors, to be replayed alongside them.

2.2.5 Institutional Stakeholders in Video Data Reuse

Support for learning and research computing in the University of Edinburgh, including multimedia services, is partly delivered on a centralised basis, and partly through staff embedded at the level of Departments or (larger) 'Schools'. Library and IT service provision has a unified structure. Various people from these support services contributed to the project through informal discussions and workshop participation, though were not formally interviewed.

The infrastructure for research computing includes high performance computing and storage facilities, and: -

- *Podcasting, streaming and webcasting* managed service.
- Digital library facilities, including the *Edinburgh Research Archive*, an open access research output repository that holds a small number of moving image items. These have been received since the University introduced in 2005 a policy that any research material examined in association with a thesis should be deposited together with the electronic version of the thesis.

The recent JISC *Datashare* project, based partly in Edinburgh, has been a catalyst for the implementation of a data repository that includes collections of the Human Geography group. The repository is at an early stage of development and does not currently include any of the group members' video corpora. During the case study period there were no specific policies in place for the curation of video or multimodal research data, with the exception of the guideline on depositing e-theses.

2.2 Diversity in the Research Roles of Video

There was wide variety in the uses of video data, even in the small number of projects the study looked at. Views on its reusability also varied between the researchers involved, their projects and research groups, the methodologies used and the purpose of the research. The study consciously sought a variety of uses in order to get a broad picture of them and, to identify patterns in their views and requirements for curation and reuse.

Some patterns were evident; sharing was heavily constrained by ethical issues around working with 'human subjects' (discussed further in the next section). Where social research projects were publishing video data on the web, this was to demonstrate the research process and its results to their research community, with potential collaboration as a general aim, rather than with the intention of offering up collections for reuse.

The three projects closer to linguistics research communities (AMI, DRS and to a lesser extent DIEM) were expressly funded to provide analysis capabilities to other domains. Not surprisingly therefore they were catering specifically for reuse, for example to allow (re)users to select video and other data sources to view in synchrony, and to enable researchers with

differing perspectives to annotate a common corpus of material. In the case of AMI, this exemplified a long standing research strategy of the Human Communications Research Centre at Edinburgh:

> *"They were interested in having a lot of different linguistics students and psychologists look at the same conversational data from different angles because they thought if everybody looks at the same data then they might learn more than if they just go out to collect their own data. So some of them would be studying the relationships between questions and answers in conversation and … some of them might be interested in how effective people are if they do a task together as opposed to doing it separately." (Senior Research Fellow, int. 30)*

DRS similarly had been conceived as an inter-disciplinary tool as well as one oriented to 'multi-modal' data, being the outcome of three 'driver' projects in Nottingham, variously involving ethnographers, linguists and psychologists. A key difference between the two was in the means of access to multimodal corpora. While both provide web access mechanisms, in AMI this is via a public web server while DRS is conceived as a client application for working within a closed network; a choice which researchers attributed to the complex ethical limitations on sharing 'naturalistic' video data.

Technology choices seemed to be more closely related to researchers' communities of interest or practice than to wider disciplinary contours; especially as researchers import methodologies from disciplines other than that which their research group is institutionally aligned with. So their *methodological* choices were reflected in their choices of tools to acquire, select, describe, store and distribute their video data. For example researchers using an ethnomethodological approach to analysis tended to follow common practices in that research community, whether they were formally aligned with geography or computing departments. The ethnomethodological aversion to elaborate coding frameworks, and affinity for informal collaborative analysis, was reflected in the choice of tools for video transcription and annotation or 'logging'. While all were aware of specialist tools for these activities, there was *mostly* avoided in favour of text editing or word processing software, and manually aligning notes or transcripts when replaying the relevant clips.

> *Logging the actual video and what's interesting, I again just use text files. I know there are applications… that let you do it, I know that they exist. But in the end text is viewable by everyone" (Computing Research Assistant, int. 15)*
>
> *We might just take handwritten notes while the video's running and then use it to go back and remind ourselves about what happens at particular points rather than annotating it and marking it up specifically. (Computing Research Fellow, int. 11)*
>
> *The change was the video editing projects… part of it is that I should become competent in a degree of video editing! So then I've, sort of picked up, Final Cut Pro and thought a little bit more about the possibilities. (Geography Senior Research Fellow, int. 14).*

The third quote above illustrates the exception. Researchers (again those sharing a similar analytic approach) would use tools suited to their research topic or application domain. For example researchers studying video editing practices in *Assembling the Line* were also using video editing tools for logging purposes, to gain familiarity with the practices they were observing. Those studying social interaction mediated by technology saw benefits in using specialist tools (such as DRS) to synchronise logs of those interactions with video recordings.

Format choices also differed (unsurprisingly) with research purpose and methodology rather than formal disciplinary affiliation. High Definition was favoured where there was a research focus on gestures or close interaction with physical objects, since this provided greater clarity, although the results from relatively high compression MPEG-4 AVC codecs provided a reasonable trade off between clarity and file size.

In the eye movement project DIEM, format choices depended on the stage of analysis. In the initial exploratory stages, researchers visually inspected the data - video clips overlayed with visualisations of their research subjects' eye movement behaviour. Again MPEG-4 codecs produced good-enough results, although the lower compression XVID codec (MPEG-4 Part 2) was used to retain enough information for the image analysis algorithms used to produce visualisations. In later experiments more precisely quantifiable results would need to be obtained from image analysis algorithms, requiring lossless compression so that as much information as possible was retained from the original video image.

Archiving practices also appeared to reflect the research purpose as much as the extent to which there was funding body interest in archiving. Where geographers and psychologists retained original recordings and occasionally revisited them, those in computing tended to be less concerned with managing video for later access, whether for their own or others' use. This was a natural reflection of the subject material; there being less of an intuitive case for computing researchers to revisit recordings of people interacting with old technologies.

> *"I think it's such an unusual situation for us. Because as computer scientists, we're always designing and building new things. So that means new data, and we're less concerned with looking over old material. Unfortunately, I think that's just the nature of our discipline " (Computing Research Assistant, int. 15).*

There was nevertheless some interest among computing researchers in building on the experience of uploading selected clips to the SEDIT website after their discussion in data sessions. The usage scenario for a more extensive archive would be to review and compare studies carried out across different kinds of settings, for example public meeting places or sports settings. The geographers already had commitments to archiving and in one case the video corpus resulting from a project – *Habitable Cars* – was already being reused by other researchers. For geographers and eye movement researchers the case for web accessible archives was similar - to promote the original research and contribute to the development of practices for analysing the moving image.

2.2 Uncertainties in Data Management Planning

The DCC Curation Lifecycle (Figure 2.1) was used in discussions of data management planning with researchers. The lifecycle model (Higgins, 2008) is intended as an aid to planning the steps needed for successful data curation at the outset of research projects. The steps are included in Appendix 1.

The researchers involved in this study were well aware of the need for planning, and had successfully submitted outline plans according to ESRC and AHRC expectations. However putting detailed plans into practice was fraught with uncertainties and concerns about sharing data for reuse. In some cases the complexities and resource-demands of publishing video data meant that project commitments to make archived video available online had not been realised.

Planning to curate video/ multimodal data is contingent on 'known unknowns' - in the data that may be acquired, what property rights and consents will be obtainable and applicable, the cost implications of capture, format and storage choices, the research community needs for context information, the technology options for data sharing, and the feasibility of long-term curation and preservation. Researchers were not hostile to data management planning, but they did have doubts about its feasibility for video data:-

> *"…it would be nice to have some kind of guidelines on these… a sort of framework within which to deal with stuff. To say 'right at the beginning I need to think about storage'. But I think things often progress in a very organic way, because you don't know." (Research Assistant, int. 15)*

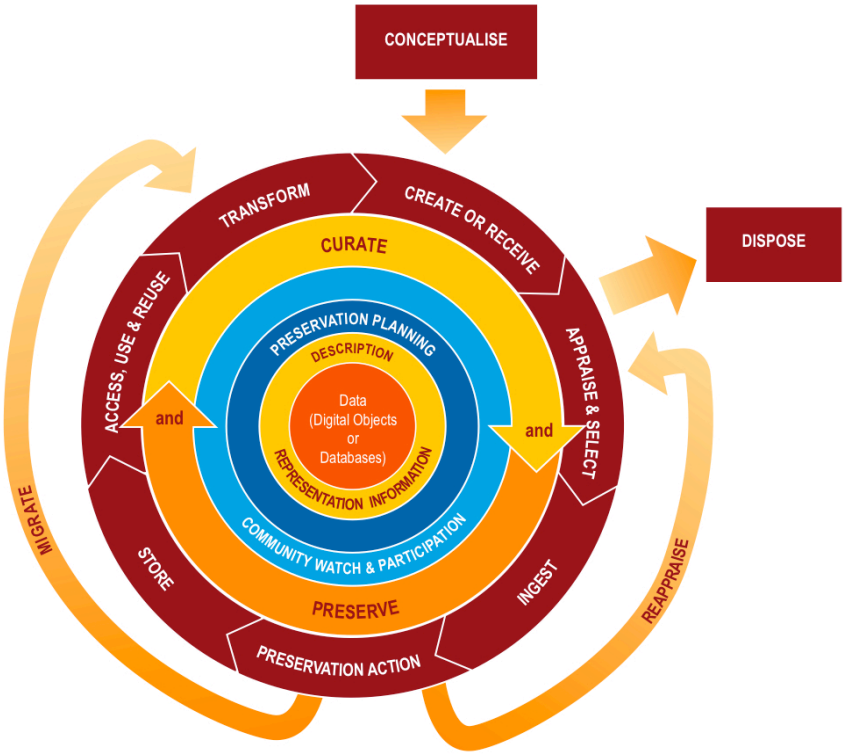


Figure 2.1 DCC Curation Lifecycle Model

As another researcher in the early stages of a video-based project put it:

> *"I've been thinking about each of these stages in a very amorphous way since we started conceptualising this project and I have to say there are so many open questions at each stage that really we don't know how to find the answers to them." (Research Fellow, int. 21)*

Some of these open questions and sources of uncertainty are identified below. These created gaps between initial plans and the reality of the quantity and nature of the video data that was obtainable, rather than insurmountable barriers to achieving research results.

2.2.1 Legal and Ethical Issues

Legal and ethical issues affect video data more acutely, although they fall into similar categories as for other media; those associated with gathering data and those with making it available for reuse, the distinction also known as 'rights in' and 'rights out'. In both cases the main issues surround rights and responsibilities to privacy and property.

Dealing with 'informed consent' in the field

Data protection legislation (the Data Protection Act 1998) applies to video as to any other recorded data, i.e. video may be 'personal data' to the extent that the individuals recorded in

it are identifiable, and treated according to the relevant data protection principles[37]. The guidelines require researchers to *inform* participants about how any personal data collected about them will be used, stored, processed, transferred and destroyed, and normally obtain their written consent[38]. This is a particular concern in observational studies, where there were uncertainties about the extent to which consent was required when recording people in public places and workplaces, and also how practically to inform people who may be identifiable but drift in and out of the scene.

NHS Research Ethics Committees' strict oversight of research in medical settings contributed to some researchers taking a risk-averse approach in applications to them, and routinely destroying sensitive personal data – including video - at the end of the research project.

Difficulties of anonymisation

Data that is anonymised is not considered 'personal data'- although it may be up to the point that it is anonymised and any identifiable sound or images destroyed. The difficulty here is that while data in text form may be edited or coded, and various techniques can be used to disguise 'micro-level' statistical data, video (and any audio recording of voices) can rarely be completely anonymised, Researchers mentioned techniques such as careful editing of facial shots, or spoken references to identifiable people or settings. Image manipulation techniques such as pixilation or blurring of faces were also mentioned. However researchers mostly considered these inappropriate for interaction research because of the loss of information on facial and hand gestures. One technique mentioned more approvingly was to *reverse the polarity* of the video image (so that it appears like a photographic negative), as this was believed to render individuals unrecognisable but preserve the detail of their actions. This was only seen as a viable strategy for presenting data however, as any shared video could easily be reversed to its original form.

Informed Consent and Risks to Reputation

Obtaining informed consent *before* recording video for ethnographic research was thought to be a 'necessary but insufficient' condition for protecting research subjects, since they could not be presumed to know what they were consenting to until after the recording had been made. Sharing recordings was considered to have potentially negative consequences for peoples' reputations. Researchers generally approached this by seeking additional consent to share recordings *after* they were made. However this solution carries its own ethical and legal problems; if consent cannot be obtained immediately after the recording this raises issues about contacting people after the recording. This may not be possible; especially in some medical research contexts where the subjects' contact details are not available.

Occasionally researchers felt that consent was not enough; they had made video recordings only after some time building up enough trust with their research subjects, trust that they felt could be breached if recordings were shared beyond known colleagues:

> *"…even though they say 'I am happy for it to be available to other researchers', people probably have this lack of understanding of how far it could be made available"*
> (Research Assistant, int. 15)

The ethical questions surrounding video were most critical for those in ESRC and AHRC funded projects using video as their main source of data, and for computing researchers in medically related research. Others in computing using video in workplace settings had not experienced significant difficulty in getting informed consent to record.

---

[37] There are various 'research exemptions' to the DPA principles, see *DCC Data Protection Briefing* available at: www.dcc.ac.uk

[38] See UK Data Archive guidelines on Data Protection at: http://www.data-archive.ac.uk/sharing/DPA.asp (July 19 2009)

Property Rights

The prospect of sharing video on public websites raised issues in the area of property rights. This mainly affected rights in data to be collected:

- '*Fair use' limitations*: Performance and copyrights are an issue in observational studies. Copyright material such as music or advertising often appears incidentally during recording in field settings, and it may be impossible to separate out the material by editing it from the activity the researcher is interested in.
- *Copyright and licensing conditions:* some researchers were unaware of the constraints placed by license conditions on material being used in a derivative work or made openly accessible on the web, and how these constraints override the 'fair use' (or 'fair dealing') usage of copyright materials for educational purposes.

One of the projects also had experience managing 'rights out', i.e. licensing the data resulting from their work; AMI Meeting Corpus applies a Creative Commons 'share-alike' license to video clips and annotations published on their site.

2.2.2 Data Collection and Development

In the ethnographic studies video would be recorded using camcorders in standard or high-definition mode, edited, logged and annotated to varying degrees of granularity; and linked to notes describing the subjects/participants and, to varying degrees of detail, what the content consisted of. However the expected volumes and types of data depended on a wide range of factors that are not known in detail at the project planning stage. For the ethnographers these would include the vagaries of recording in public places and domestic settings (homes, cars, housing estates, cafes, sports grounds). For those recording interactions with computers or other devices it would depend on whether or not this was feasible and appropriate at the time, given privacy sensitivities for example, as well as on the accessibility of settings and participants - an issue in any qualitative field research.

For the experimental researchers in *eye movement studies* the volumes and types were definable within broad limits. Video clips of various identifiable kinds would be collected and used with a (specified) eye tracker to gather a (defined) number of experimental subjects' eye movements, resulting in data volumes that could be estimated, but only broadly.

This too was exploratory research; its early 'pilot' phases aiming to narrow down a range of moving image genres and characteristics to those that would be most productive in terms of enabling hypothesis testing.

> *"…if we had a hypothesis about how people's attention may be attracted by motion in the world then we can create videos that have varying degrees of motion in them but try to keep them as realistic as possible so we're.. this pilot project we're doing right now is exploring the potential for using this kind of stimuli" (Research Fellow, int. 21)*

This depended partly on the feasibility of controlling and measuring the kinds of movement depicted in the images used as experimental stimuli. The types of image ranges from cookery demonstrations to CCTV images; diverse in terms of length, colour depth and therefore size and storage volume.

Data quality

As in all human subjects research, the quality or *provenance* of the data gathering in terms of ethical and legal factors is potentially subject to scrutiny by ethics committees. Ethnographic and eye movement researchers therefore maintain auditable records of their subjects' consent. Eye movement researchers reported that psychology journals increasingly expect submitting authors to warrant that ethical procedures have been followed.

In ethnographic studies there was no evidence of any formalised *quality assurance* process being applied to video at the point of collection. Informally however ethnographers shared their data soon after capture in data sessions, either with colleagues directly working on their

project or in a wider group (SEDIT in this case). Data sessions gave researchers an opportunity to test out with peers their interpretations, as further discussed below.

Data sessions are a form of peer assessment of the detail and plausibility of the *analysis* and its relevance to broader research themes. Sessions involve replaying the clips that are presented, so that those viewing can query the presenter on the link between whatever is (interactionally) going on in the video and whatever features of that the researcher has highlighted in transcription or talked about as an (often tentative) analysis..

Indirectly, data sessions may be considered a form of data quality assurance, but it is not one that lends itself to measurable criteria. Researchers did not use the term, or apply any explicit quality criteria to their data. Data sessions nevertheless depended on the *interpretability* of the clips– the extent to which session participants could make sense of the interaction portrayed. Selection therefore tends to favour sequences of visible and audible interaction that are played out in no more than 5 minutes, since clips are viewed and then discussed, making longer sequences more problematic. The *clarity* of the clips is also a factor in their utility as evidence. Especially valued are sequences where the meaning of a word or phrase has some *unexpected* relationship with the context that can be demonstrated through repeated replay, supplemented by any further details of the setting that the presenter can provide.

For the eye movement researchers data quality was more measurable and predictable early in the *DIEM* project. These were aspects of the eye-tracker instrumentation and its capabilities, measurable in terms of *precision* and *accuracy*. Eyetracker *precision* refers to the sampling frequency, in this case an Eyelink 1000 manufactured by SR Research samples at 2000Hz, i.e. providing 2000 measurements of the position of each eye per second. *Accuracy* is a statistical measure of confidence limits in those readings, identified by calibrating the eye tracker using a vendor-specific procedure. This is done before and after each subject is shown the video (experimental stimulus); each person is asked to focus on a simple stimulus such as an image of the letter X to centre the camera on a reference point for subsequent eye movements.

Format troubles

Format choices also impact on later storage options and there are few standards for the more esoteric data gathered in multimodal studies, including eye-tracking data. There are on the other hand many well-defined 'standard' file formats for video data, and several for managing video metadata. The issue for researchers in the case study was not lack of standards for video data but too many of them. This was compounded by a lack of readily available information on platforms for managing multimodal data.

The advice available on digital video formats from *JISC Digital Media* pointedly notes that "Choosing a file format can prove overwhelming for someone new to the world of digital video" but that the key elements are the codec (compression-decompression algorithm) and the 'wrapper' file binding the separate results of audio and video compression[39]. A source of confusion is the relationship between *file formats* such as .mp4 *compression standards* (such as MPEG-4), *codecs* compliant with MPEG-4 such as XVID or DIVX, and media player software such as *Video for Windows* (e.g. DIVX is a codec, file format and media player). It is difficult to unravel the interdependencies between these, the research aims and methods, and the choices of equipment and data management software, which may support only a limited number of formats.

Researchers interviewed from the ethnographic and the eye movement projects talked of the difficulty of selecting and specifying equipment at the project planning stage, as this depends on finding an effective trade off between the equipment their research budget can afford, the formats offered, the storage requirements of different levels of compression, and the current

---

[39] Choosing a digital file type: available at: http://www.jiscdigitalmedia.ac.uk/movingimages/advice/choosing-a-digital-video-file-type/

standards of the day.

> *"When I did early work I had to think… I didn't know quite what the plan would be, because I was still trying to learn what were good formats to work in and so there's lots of craft skills to pick up there, so planning at that stage was really, would have been… you could have done it and I guess I did do it, but they were much more likely to miss the mark because you hadn't yet learned… I settled on a particular encoding format that seemed to be pretty good space wise, but wasn't too much of production level quality, and once having tried out, because also it takes a lot of mucking around and experimentation to compare different ways of compressing or encoding video to see which seems to have the right trade off for you and… so that was, sort of, settled and that did work well for a lot of projects."* (Senior Research Fellow, int. 14)

Several factors have recently made this more difficult: firstly the lack of widely supported standard formats that meet the criteria for an *archival* format (lossless compression, non-proprietary methods), and, secondly, the added need for archival storage with the advent of hard-disk camcorders.

The significance of these is that for the 10 years or so that mini-DV tapes have been a de-facto standard for storing the data captured from camera they have in effect been used by researchers as the media for their personal research archive. Hard-disk based camcorders have therefore made archival issues more especially acute:

> *There's been a switch up in terms of using HD for the first time, so that's raised a whole series of issues about rethinking what sort of compression is going to be done and simply how much footage I then start trying to collect, because it becomes like the beginning of using video, you're running out of drive space very quickly, so you shouldn't really collect too much material, and rather than compress it all down to a small format, because I wanted to see a lot of what happens on computer screens, leave in that quite detailed… which means you can get too compressed and because the ideal in video editing is also not to use overly compressed footage, because it's a bit hard for the computer to edit because it takes a lot of processing power to work with compressed footage, because the computer has to do all sorts of rearrangements to then get it into a, sort of, frame by frame basis.* (Senior Research Fellow, int. 14)

Lack of metadata standards

Metadata standards, as Chapter 1 described, commonly refer to *descriptive* metadata (including any classification or controlled vocabulary), *technical* metadata (describing technical processes used to render the object), *administrative* metadata (including contextual details of the object's creation and version control), and *preservation* metadata (including details of data integrity checks, or actions taken to migrate data to new formats).

Researchers were not using any *descriptive* standards. There are no widely accepted descriptive metadata standards for ethnographic materials in interaction research, or of eye movement experiments. There are some potentially adaptable contenders, emanating from linguistics domains that we return to in the next chapter.

Descriptive information was recorded as 'logging' data; the key words and phrases used to tag particular frames or sequences. As already noted, logging is a term from video editing rather than research, but overlaps with 'annotation' and 'coding' although the latter are more usually associated with using controlled vocabularies or developing theoretical frameworks.

The *technical* metadata needed to render video is normally automatically recorded by the encoding software and stored in whichever wrapper file format it supports (e.g. Quicktime, AVI). However available metadata standards for managing descriptive annotations and for *packaging* them with technical information, such as MPEG7/21 and METS, are not widely supported in commercial software for managing video data.

Researchers used a variety of software packages to log their clips and manage the technical

metadata generated at the point of capture. Problems were experienced in the transition from working individually on clips to working collaboratively. One project in particular had difficulty finding software that would allow individuals to easily log their clips, while also providing common metadata administration functions to, for example, edit terms applied to collections of files at different level:

> *"There was thousands of files and there was just … maybe some ten, twenty percent I could do that work adding an exact description of what the file consists of, of what is there, who is the author, and what the title and key words are. Some of this stuff was coming with the file automatically … and I found that quite useful, but the problem then was how would you transfer in between different files this metadata…"* (Research Assistant, int. 6)

Ethnographic researchers used notes to record contextual details of the setting being recorded, although these notes were seen as for personal interpretation,:

> *It's not driven in the same way …by the notes, it's much more driven by the video footage and the transcripts, with the notes being a, sort of, fall back for some sort of context, particularly in work settings, for that material* (Senior Research Fellow, int. 5)

The eye movement researchers were considering options for metadata but, in the absence of any directly relevant standard, the experimental details would be recorded using the vendor's proprietary software for eye movement experiments, *Experiment Builder*. Integrated with the eye tracker, the software provides a drag and drop interface for designing and executing an experimental workflow, enabling common tasks such as calibration to be performed by inserting an object in the appropriate place, and also gathering the results data. The workflows and results help to maintain an experimental record, although with an obviously high degree of vendor dependence.

The other sources of documentation for eye tracking experiments are the *ethics documentation* outlining the methods proposed for data gathering and experimentation, the individual subjects' *consent forms and demographic data,* and the *lab notes*. These notes do not attempt to record each detail of the experiment, but exceptions to the plan:

> *"…it is really just to record things that have gone wrong - for example if someone coughs or sneezes - or if someone walks in and distracts attention - there will be blips in the data that are not otherwise explainable so these get recorded but they are not common" (*Research Assistant, int. 17*)*

2.2.3 Short-term Storage and Management

Since multimodal data, video especially, consumes so much space yet is easy to capture the most significant issue for many of the researchers was the trade-off between the costs of managing storage and what can *usefully* be captured.

> *"The space issue I think is the biggest problem. And even if you want to share a video, it's massive. It's still massive. And the other thing is that the problem is not going away, because as storage gets cheaper, qualities get higher. … I think I've just always done it in a really piecemeal way, so just expand and keep expanding the storage space. Start spreading storage across awkward volumes. … I think it's always going to be a problem, especially I don't think it's possible to solve or even plan for…I suppose you can plan to some extent for massive amounts of space. But as I said, it's always going to be more than you think, essentially. And I don't think it's going to go away either."* (Computing Research Assistant, int.15)

Storage management was viewed as a perennial problem that is inherently difficult to contend with. For those projects making substantial use of video it is inter-related with all other stages in the curation lifecycle. The three main reasons for this, and the uncertainties surrounding it were the exploratory nature of the methodologies used, the sheer size of video without lossless compression, and the logistical and ethical complications of planning for

shared access to whatever parts of the data are retained.

Data reduction in exploratory projects

Multimodal data is obviously 'data' in the sense of being digital information; a more open question is the point at which it acquires a role as *research* data, i.e. evidence in the research process. Much of the data gathered will not be used, but selecting the relevant material depends on having adequate short-term storage.

Data reduction has traditionally been a concern for those, such as ethnographers, who predominantly use qualitative data analysis methods but to a far greater extent where video is concerned. Researchers interviewed typically used three main stages of data reduction along the research cycle in Figure 1.3:

1. A first-pass review of the material captured from the camcorder, to log selected sequences, i.e. record descriptive key words and phrases against their time-codes. Around 10% of the captured material is kept at this stage.
2. Selection of relevant clips for transcription and analysis in 'data sessions'. Selected material might also be used in teaching. Footage is reduced again to around 10% of the logged material, which may be compressed by transcoding so it may be moved around more easily.
3. A further 10%, *i.e. 0.1% of the original footage* would be analysed further for publication. This may involve referring back to the original material, and encoding more of the sequence, or re-encoding it for greater clarity.

An unresolved question that arose in the case study workshop was whether the 90% of recorded footage that is *not selected* at the first stage is actually 'research data', or might be better thought of as 'deferred observations' until such time as it is revisited and selected or discarded. This would have implications for the role of support services in curating it.

Eye movement studies using moving image stimuli are also exploratory, meaning that some of the data (stimuli and results) would not be relevant to more detailed experiments leading to publication. As this was yet to happen in the DIEM project the amount of data tat would be selected for publication was uncertain, except where there were known copyright issues with the source material that would prevent it.

2.2.4 Access, Sharing and Reuse

The legal and ethical uncertainties mentioned earlier are such that there is limited public sharing and reuse of naturalistic video data in the projects featured in this study (or generally as far as researchers were aware). Multimodal corpora are designed for reuse, but in the researchers' experience this has mostly been for repeated analyses *within* a research group. Managing IPR risks was a particular concern of the eye movement researchers aiming to construct a public archive, given that they mainly used clips from 3rd party sources.

Among the ethnographic researchers, some participants talked of the difficulty replicating in digital form the nuanced permissions that normally govern the informal sharing of data. Even among the informal networks of researchers involved in data sessions, the norm is that clips are viewed together but not subsequently shared for reuse beyond the core research team:

> *"I think that's what's important about data sessions, is that you watch it, you sit and watch it. It's like watching a film isn't it. You go and watch a film, but you don't take it away with you. But you do have some kind of shared experience. And it's nice, because it means that other people can see your data and can remember it. And you perhaps use it to inspire some other thing they do or whatever, perhaps. But they don't have a record of it, which is the issue, the record* (Research Assistant, int. 15)

Despite this caution, some of the SEDIT ethnographers were sharing clips online after discussing them in data sessions, normally with password protection.

Public access to video data was available in the two projects using video for research on meeting support technologies. In both cases the ethical complications of using video captured 'in the field' had been circumvented. In the *AMI meeting corpus* the video recordings were of volunteers acting out scenarios. In the case of the *Branded Meeting Places* project the video subjects were the researchers and graduate students workshop participants. Clips were embedded in a blog, which was publicly available but mainly intended as a vehicle for (distributed) project participants to catch up with others work, or to help transcribe group discussions.

Access issues were difficult to separate from storage management and the logistics of handling high-resolution video. Collections varied from approx 6 hours of footage (anticipated for *DIEM*) to 240 hours (*Habitable Cars*) and 600 hours (*AMI Meeting Corpus*). Across the projects encountered in the study, the AMI project corpus was the largest publicly available collection with 100 hours of meeting event duration gathered at the time of writing. As each meeting involved separate recordings from 6 camera angles, this comprised 600 hours of footage. The size of the corpus meant web dissemination was not a viable way to meet the needs of all its users:

> *"Our major problem is that if you get our data in its entirety including the high resolution video for doing proper video processing work it's a terabyte… so we can't even serve that over the web, we only serve the needs of the non-video community from the website and even then people complain about there's a lot they have to download. But we have to ship out on firewire [disks] from a site in the Czech Republic for the video processing people and this is a major hassle" (*Senior Researcher, int. 30)

Common data management questions across the projects included where and how to maintain access to the original footage, and how to store and link archival and low-resolution distribution copies. The main issues and uncertainties here were:-

- *Estimating and budgeting for storage needs.* Research groups in Human Geography more used to managing their own data were inexperienced in planning for server and backup resources at departmental level.
- *Managing external hard drives*. The availability of cheap high-capacity external drives was highly valued, to cope with unanticipated storage needs and for their transportability. These were thought cheaper and more convenient than centrally managed backup facilities, although researchers were also aware they lacked information on drive reliability, and their proliferation might lead to difficulty finding resources.
- *Local computing support* was a key factor; researchers valued it in the areas of web video streaming and server management, whether it was present and appreciated or absent and missed.
- *Participant consent issues* were a potential factor because participants' rights under the Data Protection legislation to withdraw at any time from a project could entail finding and editing them out of any footage stored.
- *Participatory research* in Human Geography projects were being initiated that aimed to directly involve research participants in producing ethnographic research materials including video. As well as the more obvious ethical, IPR, and access management issues this raised new issues in managing short-term storage.

2.2.5 Deposit and long-term preservation

The post-project fate of their data was inevitably an issue for researchers, who mostly did not have any mandated data repository. Without further funding data collections would be abandoned unless, as one researcher put it, they could be "chopped up into Masters-sized chunks". Institutional data services being at an embryonic stage of development, the main issues identified here were:

- Facilitating progressively wider *data exchange*, to track how others have used

researchers' own data and, in return, accessing and annotating others' data.

- *Federating curation resources* so that data can be effectively curated locally, whether to help researchers prepare for deposit, or to provide a local repository integrated with services at the institutional and national levels.

Underlying both points is that researchers were interested in keeping their datasets 'alive' in the long-term; and, those with intentions to publish data saw its reuse by others as a means to do that.

Curation as exchange

One senior researcher experienced in working with video identified curation with the period of a project when the data has been pulled together and organised to the point it can be 'handed over' to someone else:

> *"The first time you ever get a project, the funding body are generally saying, you must also think about also curation at the end, set aside time for it, but nevertheless… it's [more] often that moment of hand over where you pull together a lot of the material… hopefully, you've been organising it… and I guess the more I've done this kind of work, the more I've organised the data along the way, where there's thought given to the fact that it will be handed over…."* (Senior Research Fellow, int. 14*)*

A key point here is that data may be re-organised, for example to pull together clips and re-classify them according to the themes of a publication, conference or data session.

> *"…that's the thing, when it's live still, there's often, you're willing to, sort of, keep… trying to move things around which may fit better under a different organisational format"* (ibid.)

This emphasis on curated material being always incomplete was also found in the AMI project. The need here was to facilitate the exchange of annotated resources, through better infrastructure to support licensing initiatives like Creative Commons:

> *"I want to mention something else which is that the data set isn't ever finished in a sense. We're not going to add more signal to it, but we've released it under a Creative Commons license that's got a Share Alike component to it so that we've released all these annotations which are valuable things but if other people annotate and they want to share them with other people then they need to share them back with us, and we're trying to encourage a community feeling for that. So there's a problem with life span for this kind of data because we've set up a maintenance problem for ourselves for a while. You can put the Share Alike licensing on it but it has no teeth. It's too difficult for people to do anything, even if they want to, unless we also keep a repository where people can contribute data back in and then we do new data builds and distribute a new public release that's got the extra stuff in it. I have no idea how we fund that in the long term."* (Senior Research Fellow, int. 30)

Federation of resources- the 'productive repository'

For participants in the case study workshop – researchers and data support providers - the principle of devolving curation to local level (school or department) repositories for video data had general appeal. Potential benefits were that the volume of material that can practically be curated at this level is self-limited to what the originating group can work with and readily understand the context of, providing realistic boundaries for their curation effort. Centralised institutional services would be needed to manage storage, and provide services such as video streaming and format migration (transcoding) for resources of identified value.

The Human Geography group, for example, was considering implementing a 'productive repository' with the intention of integrating it with the institutional data repository. The most salient need here was to support the collaborative aspects of video research, to support intensive logging and tagging. This could then be the basis for an online resource that could

be browsed and annotated, with those annotations becoming part of the curated object (as discussed at the end of chapter 1).

For long-term curation there was a need (beyond video data) for an institutional model, which might be seen as a form of 'client-server' approach to data repositories. Something like a 'research executor' role was needed for collections of data whose owners had departed the institution, dependent as that would be on access to context information.

2.3 Data Sessions and Contextual Information

Creating contextual metadata is one of the most significant cost elements in data curation, and it is often stated that overall costs can be minimised by creating it early in the research process (e.g. Beagrie et al, 2008). An important aspect of SCARP is to understand how to embed metadata creation in the research process in ways that are of immediate benefit to researchers or at least do not burden them. The case study used participant observation in *data sessions* to better understand whether and how and whether they afford opportunities to do this. SEDIT data sessions were attended over a period of 5 months. Three sessions were video recorded, and field notes were made on other occasions. Segments of two of the recorded sessions were selected for transcription and analysis.

2.3.1 Producing and Managing 'Context'

SEDIT data sessions mostly involve participants who are not part of the research team gathering the data, so they can be considered a form of data sharing if not secondary analysis. Researchers wanting to analyse social interaction data created by others need to know something of its context, as do researchers in any domain. What constitutes that context and how to represent it is a matter of debate that is far from settled in qualitative social research generally and, as Chapter 1 pointed out, especially difficult when the research is on contextual aspects of interaction, recorded in rich media.

The main question to be explored was; *how is 'contextual information' worked with in data sessions*, i.e. what is it taken to be and what is accomplished with it? Chapter 1 partly addressed the first part of that question. That is, for interaction researchers, the ongoing flow of talk and action in the recording is as much part of the context for analysis as is any description provided of the original setting that the recording was made in. So from that perspective in the e-Social Science project *MiMeg* the activities of annotating video and providing descriptions of it were treated as 'embodied collaborative actions', to be supported through real-time interactive tools, rather than as the creation of records for asynchronous use (Fraser et al, 2006).

The studies of data sessions carried out in *MiMeg* describe how researchers presenting their data begin by providing some background information about the nature of the scene displayed on screen, initially as a still image. This was seen to 'establish a common ground' for the subsequent discussion, and would typically refer to who the videoed participants were and what they were in the course of doing (ibid.)

For that reason, in the case study data sessions were initially approached on the assumption that characterising how researchers *introduced* their data would help understand the scope and depth of description that would be expected of context metadata. It would at least identify how much researchers want to know about the setting before commenting.

From some perspectives, particularly those closest to Conversation Analysis, transcripts need very little introduction. One interviewee put it as follows:

> *"I get the feeling there's only a few specific sections of qualitative researchers who … think that the data is of interest in itself regardless of the project it was created out of or the sort of social context of that particular interview or event which is, of course, people who do basic conversational analysis type work, where they're quite happy to look at a transcript and know nothing beyond some one sentence description at the*

> *beginning of what it was and the less context they have, the better, because they just want to get at what the inherent internal organisational features of it are."* (Senior Researcher, int. 14)

Another factor favouring introductions that are less laden with contextual information is that participants in SEDIT data sessions are mostly regular attenders. As such they generally have some awareness of each other's current projects and interests, and the kinds of contribution typically made at data sessions. On the other hand SEDIT sessions are *inter-disciplinary* gatherings, and although some participants have a conversation analytic background, others have more of an interest in being familiar with the setting. And while most will know generally what topics the other attendees are working on, they cannot be assumed to have the shared knowledge that researchers directly engaged in their project would have.

The presenter and topic of a SEDIT data session is announced to regular attendees by email from one of the organisers. Who is to present and on what is normally arranged at least a few weeks in advance. Presenters may be visiting researchers, but are often drawn from the regular attendees. Through the informal discussion that accompanies data sessions (itself one of the reasons for attending) organisers are generally aware of who is working on what data and may want to present.

Data session participants therefore generally arrive already equipped with some very general 'context information'. E-mail announcements are typically sparse on detail, but range from one or two words on the general topic to several paragraphs, accompanied by a full transcript. Presenters are anyway expected to provide transcripts at the start of a session. In the data sessions observed, presenters also gave some account of the setting portrayed in their clip. They highlighted or were asked by the participants (we can refer to them as 'viewers') to describe the subjects' reasons for being there, their roles in a workplace, how and why the video was obtained and so on.

It may seem obvious that this context or 'setting description' as we might call it, fits within the scope of contextual information or metadata as defined in standards such as OAIS. It might be argued therefore that this point in the research process – the preparation of selected clips for sharing in a data session - is an appropriate point 'upstream' from the archival ingest process to record metadata on what the video clips are about and why they were recorded.

Observations from participating in these sessions indicate otherwise. Although presenters should be able to describe the setting of their recordings, sharing this as structured 'context information' is neither necessary nor helpful for 'data sharing' at this stage. This is not because participants generally already arrive with some informally gleaned awareness from previous conservations about the presenter's current work. They are just as likely to know no more than was revealed in the meeting announcement. The main reason contextual description is not required before data sessions is that an important part of their purpose is to clarify what the *relevant* 'setting description' is, and establish what bearing it may have (if any) on how interaction between the videoed subjects is organised.

2.3.2 An Example: Video Editing

A brief example demonstrates how 'context information' is used in data sessions. In the session, the presenter ('Frank') is aiming to analyse what is happening in a clip of two people editing a documentary video.

1.	Wilma:	But this is her being shown it as it is (.) the most recent version of the
2.		edit as it were
3.	Frank:	<yeah (1.0) > and with her orientation to working on it together
4.	Wilma:	Yeah↓
5.	Frank:	So she's not (.) like (.) > one of the ways it would be important to
6.		differentiate is < she's not like in a critics session where he shows it and
7.		gets all these remarks back from people [and then leaves to fix it himself
8.	Wilma:	[yeah
9.	Frank:	She's beside him and they're working through it together sort of step by
10.		step sequence by sequence and ehm she's not got her hands on >that's
11.		what's interesting as far as it works out < she's not got her hands on the
12.		controls but I think they really are trying to do it collaboratively (.) ehm
13.	Wilma:	Hmmm
14.	(3.0)	
15.	Frank:	> and part of how it's played out is ↓interesting
16.	Wilma:	Yeah its an interesting thing just in terms of (1.0) the fact (.) yeah it is
17.		supposed to be a joint effort ↓ (0.5) but (hhh)

Transcript extract (see Appendix 3 for transcription conventions)

As is normally the case the presenter introduces the session with a pre-selected clip, one of several of around 2 to 5 minutes to be discussed in a session, paused at the beginning and projected onto a screen. In this case the still image shows a man and a woman who sit with their backs to the camera, in front of a desk on which there are several computer monitors and some other equipment. In the first 20 minutes of the session Frank offers a description of what the clips show – that they involve the man and woman editing footage for a documentary for the community organisation they work for, that both have editing experience and the woman has a background in the arts, and that their work involves 'assembling materials'. This is an unusually long introduction, necessary because the three people viewing the clip have varying prior knowledge of professional video editing.

The transcript (extract 1) is from a recording of the data session, and at a point about halfway through the two-hour session. Among the things that the discussion has already established are that the man and woman are jointly working on a task that would more typically be done by an editor working alone; frame-by-frame editing of the soundtrack to replace a voice-over with some music. Also Frank has indicated that he only recorded this clip in the past week and has no firm thoughts about it, other than that the clip is unusual (for the kind of work depicted) in showing detailed screen work being discussed rather than simply performed.

This short extract from the data session has various instances of presenter Frank providing 'setting information' about the video editing shown in the clip, partly in response to questioning by viewer Wilma; for example that the kind of session being viewed is in a different category from a 'critics session' (line 6), and that the woman editor has an 'orientation to working on it together' (line 3) that is pertinent to the organisation of this particular editing session. These aspects of the 'setting' are not *observable* in the recording.

The exchanges in lines 9 to 17 by contrast are observable in the clip. Viewers can readily see, by replaying the clip if need be, that the two editors are sitting side by side and

performing a step-by-step sequence of actions, and that the male editor physically has his hands on the computer. Frank is thereby not imparting anything here that is news (or 'setting information') to the other participants, but seeking some shared interest in the interactional features of the sequence. More specifically Frank suggests that the interaction shown is organised collaboratively; that the male editor is not just responding to the female's critique, and invites the viewers to look closer at 'how it's played out' (a contextual reference to the clip that has been repeatedly played back in this session).

Wilma is noticeably willing to go along with Frank's proposal while remaining sceptical that what they are witnessing is indeed wholly collaborative. Much of the remainder of the session is concerned with looking in closer detail at a sequence of interaction which puts this question to the test, with the aid of a second clip that Frank locates, showing the sequence from an alternate camera angle. By the end of the session this clip has attained the status of evidence, demonstrating how the male editor is evidently not in control of the action despite being physically at the controls.

2.3.3. Learning in and from Data Sessions

The example discussed above is an ordinary example of interaction in data sessions, and (although not in much depth) adopts a similar approach to analysis. It is no great surprise that, in trying to make sense of what is going in a clip and taking that apart in some detail, viewers ask for and are given whatever details may be relevant at the time. Despite its ordinariness it demonstrates some points relevant to the use of 'context information'.

The first point is that, at the outset of a data session, what a clip is 'about' is in a sense up for grabs. The presenter may have a research theme that is known to the viewers; the nature of 'repair work' for example, but that can only be justified as a description of the clips being presented to the extent that the presenter justifies any claims made about the clip - by referring to specific interactions, and responding to any requests for additional information about the context that will satisfactorily embellish the viewers prior knowledge of that form of interaction. Data sessions *produce* 'context information' as an inevitable by-product of making sense of what is going on in a clip. Presenters are, in a sense, moderators in a dialogue about what kinds of description are reasonable.

Secondly, while data sessions also clearly depend on contextual information, the extent of it cannot be pre-determined. Nor do the presenters deliver all the context needed when introducing their clips, before the session proper gets underway. Presenters 'answer for the data' throughout data sessions, and they cannot do that all up-front because, however much they know of their clip and its setting, they cannot know in advance what level of detail or way of formulating it will help their audience to make more sense of what is going on in the clip than they can themselves. The 'internal organisation' of the portrayed interaction becomes apparent through the interaction between presenters and viewers.

Thirdly, presenters may downplay some aspects of a setting that they know (or have inferred), so as to avoid pre-empting the viewers' reading of what is going on. In the example above, a plausible reading of the female editor's role would be 'director'. Identifying that in advance would not in this case have helped and might have hindered the detailed examination of the interaction and the role of the editing technology in that.

Other aspects of SEDIT data sessions were apparent from observing them. It is important to note that although the analysis is interpretive, it is not the case that 'anything goes'. Participants offer their take on the unfolding events in the recording presented, framing them in terms of other kinds of talk or interaction they seem similar to – for example as 'teaching and learning', or 'manoeuvring in an enclosed space'. These suggestions are not always accepted; a presenter may – since it is his or her data to be worked on – dispute other viewers' characterisations. A viewer may see in a clip something that reminds them of something from their experience or from reported studies, but such reminiscing is only taken to be relevant if it lends itself to some description of what is going on that is demonstrably plausible by replaying the sequence in question. Nevertheless sessions are not normally

confrontational; observations are picked up on and explored in more depth or fall by the wayside, presenters are expected to argue for their analysis and warrant their claims, but destructive critiques are avoided.

Finally, data sessions are (as pointed out in the literature review) *occasions for learning*. Their format could be described as a 'peer led tutorial; led by a presenter who has some expertise in the form of interactive work shown in the recordings, and with the expectation that others present will gain something by attempting to make sense of it, and might present their own material in a future session. Graduate students are encouraged to attend and frequently do. One researcher compared data sessions with conferences in that they were – "as much about catching up with what people are doing as for any serious analysis". This suggests possibilities for combining curation needs and workflows with those for e-learning.

3. Strategies and Sources for Curating Video

The chapter draws conclusions and recommendations, first reconsidering the DCC Curation Lifecycle in light of Chapter 2. The difficulties of planning video data management at the outset of an exploratory research project are that many decisions will be uncertain and the risks and steps required will need to be revisited periodically. The chapter first proposes three phases of the typical research cycle, each of which can be aligned to an iteration of the Lifecycle Model, to give more emphasis to curation as a *phased* activity (as it is in other models). The chapter also identifies strategies and sources of information relevant to developing curation further in the two main case study areas; ethnographic video analysis and eye movement research.

3.1 Aligning the Research and Curation Lifecycles

In social research domains traditionally grounded in repeated study of collection of examples of human interaction, the addition of video complicates data management planning. As the last chapter identified, decisions made at the outset of projects like those discussed in this case study may well need revision in light of changes in the data gathered or the options available for dealing with them. Rather than fit the DCC Curation Lifecycle onto the typical research cycle in an attempt to plan both at the outset of a project, an iterative approach is likely to be needed if video data is intended to be a substantial element of the research evidence to be collected.

Phased approaches to curation are a significant feature of recent guides and models of curation; the ICPSR *Guide to Social Science Data Preparation and Archiving* (ICPSR, 2005) for example is organised around seven more or less linear stages from 'proposal writing and planning' to 'after-deposit archival activities'. While this presumes that curation activities are targeted toward deposit in a single archive, recent work in the Datashare project models curation around several stages, envisaging institutional repositories offering a curation workspace to researchers before they are ready to deposit in any domain archive. According to this model, such 'staging repositories' should be the focus of 'partnering moments' between researchers and data curation services at the institutional level (Green, 2007).

The DCC Curation Lifecycle encourages research data creators and others to consider the whole lifecycle of their data so they can manage it better whether it is primarily maintained for access by the core research team, a wider collaborative group, or for publication by an institutional or domain archive/repository. This does not apply uniquely to video or multimodal data, but the storage management and access issues that are a perennial issue with moving image data in social (and medical) science make it more important to consider curation 'in the round' wherever it is stored, and at whichever phase.

Considering the DCC Curation Cycle alongside the 'curation domain' approach and the research cycle shown previously in Figure 1.1, Table 3.1 and Figure 3.1 illustrate how the Curation Cycle across three phases of video data curation. These borrow from the '*curation continuaa*' model (Treloar and Harbroe-Ree, 2008) which envisages three curation 'domains' with differing characteristics of data and metadata. The role of curation services is this view is to *support the transformation of the data as it moves from one phase to another.*

Phase	Characteristics
Planning and Piloting	Less metadata, more items, larger objects that are often continually updated, researcher management of the items, less preservation, mostly closed access and less exposure.
Project Curation	More metadata, fewer items, smaller objects, researcher management, more preservation, and less restricted access.
Long-term Curation	More metadata, further reduction in items or size, organisational management, more preservation, more open access and exposure of metadata for harvesting.

Table 3.1 Curation phases

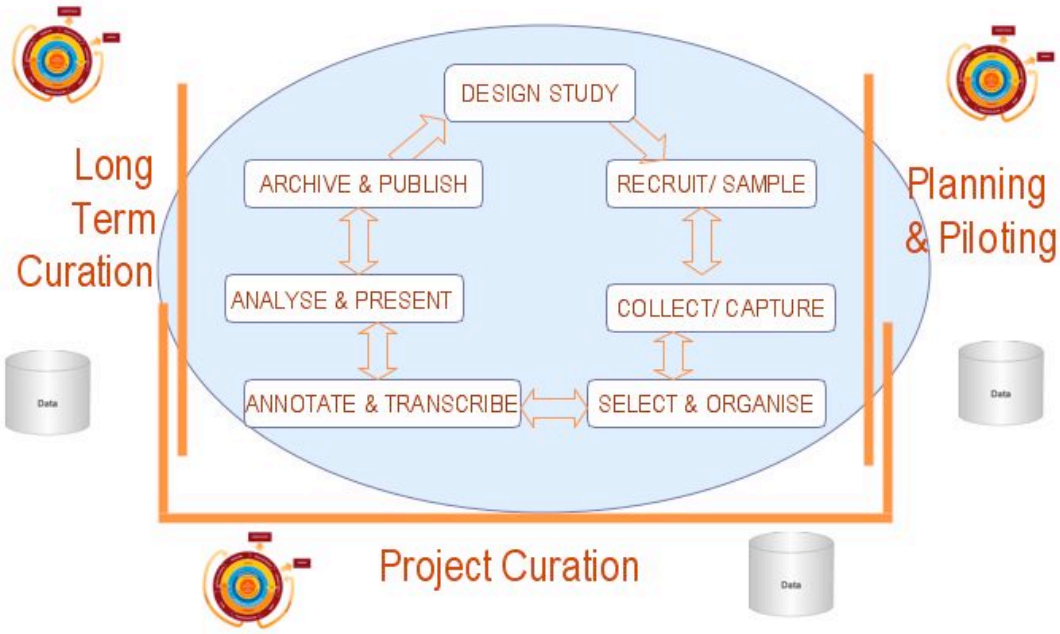


Figure 3.1 Iterating Curation Cycles across Research Phases

The overlaps in Figure 3.1 represent the periods when data is transformed to meet the needs of the next phase; each corresponding to the *transformation* step of the DCC Curation Lifecycle. That is, data is transformed to meet the needs for access, use and reuse anticipated in the next phase, based on a re-assessment of the the Data Management Plan and any risks (to data re-usability) foreseen at the beginning of the project. Figure 3.2 below illustrates the main elements of risk assessment as applied in the DRAMBORA approach for repositories.

3.2 Strategies and Sources for Curating Video Data

The remainder of the chapter describes strategies and sources for addressing the issues and risks identified, and follows the structure recommended for a *Data Management Plan*.

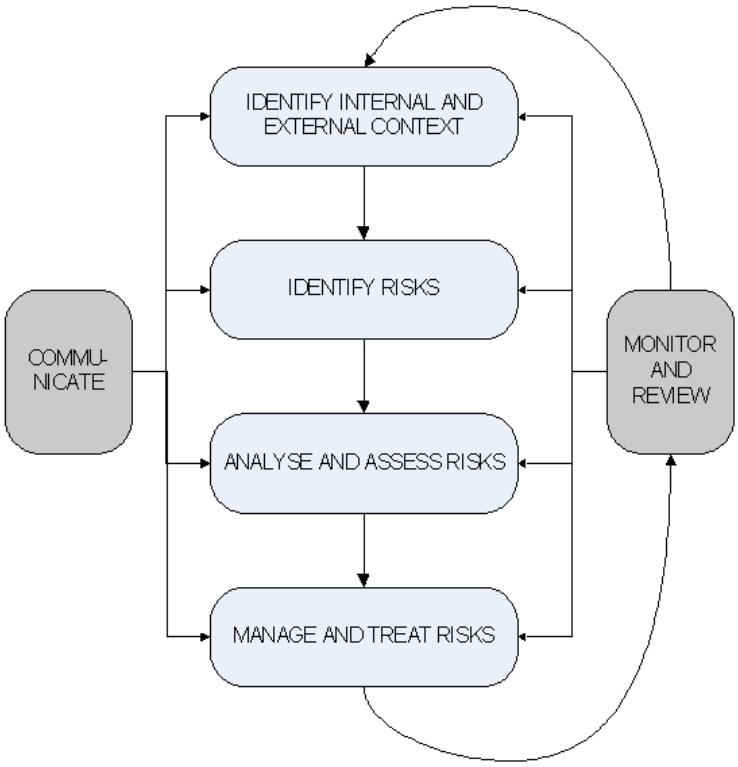


Figure 3.2 DRAMBORA Risk Assessment Process

3.2.1 Data Management Planning

DCC provides a generic template to assist in developing the Data Management Plan that research councils increasingly expect research investigators to complete as a condition of funding[40]. Data Management Plans at the research bid stage are typically short documents that will commit the bidder to follow relevant 'good practice' principles, and identify decisions made or options being considered. The plan can then be developed further, so it may be used throughout the project to assess and manage data curation risks so they do not impede the research.

3.2.2 Legal, rights and ethical issues

A Data Management Plan should identify the ethical and privacy issues affecting the research and how these will be addressed. It should also indicate how intellectual property (IP) rights are to be protected, including any 3rd-party rights in the data collected ('*rights in*') as well as rights that apply to work newly created through the research *('rights out'*).

3.2.2.1 Property Rights and Licensing

Most IPR issues affecting the use of video as research data will also arise with its use for e-learning, and general guidelines are provided by JISC Legal[41]. The risk management toolkit for IPR (Korn and Oppenheim, 2009) produced for the JISC Strategic Content Alliance (SCA) lists common risks and outlines these strategies for mitigating them:

- Use *credit lines* to show the works in which rights are owned by third parties

---

40 Data Management Plan template at: http://www.dcc.ac.uk/docs/templates/DMP_checklist.pdf (June 12 2009)

41 e.g. Casey (2006) *Intellectual Property Rights (IPR) in Networked e-Learning: A Beginners Guide for Content Developers* JISC Legal Information, available at: http://www.jisclegal.ac.uk/Portals/12/Documents/PDFs/johncasey.pdf (June 12 2009)

- Use and observe copyright statements and *Creative Commons* licences,
- Restrict access by password to certain content, e.g. of high commercial-value
- Use *indemnity clauses* in agreements and licences for materials that have permission granted by third parties, to ensure that the parties who are granting permission have secured the right to do so
- *Acceptable Use Policy*: an agreement that is physically or digitally signed or accepted to acknowledge the user is aware of the organisation's stance on copyright infringement and will exercise due diligence and care not to breach this policy.
- *Notice and Take-Down Policy*: a policy ideally linked to every page that has contributor content, highlighting the organisation's stance on copyright infringement and specifying the conditions for removing content. (ibid.)

The SCA toolkit provides sample policies along with further guidelines for any public sector organisation working with web content; they are not specific to research data or to video.

More specific examples of risks encountered in the present case study were:

- *Sharing video on 3rd party websites*: researchers may be drawn to video sharing websites that allow video to be uploaded very easily, and shared in password protected areas with known users. The benefits of this need to be traded against the risks, which in this case were of researchers inadvertently breaking the terms of their consent agreements, or granting unsecured rights (see below). The *Vimeo* web video sharing platform for example operated, at time of writing, under the following terms and conditions:

  > *"VIMEO shall be entitled to unrestricted use of any Submission for any purpose whatsoever, commercial or otherwise, without compensation to the submitter. You also grant each user of the VIMEO Site and Services a non-exclusive license to access your Submission through the VIMEO Site or any other access point authorized by VIMEO, and to use, copy, transmit or otherwise distribute, perform, publicly perform, create derivative works of, and display your Submission to the extent expressly authorized by VIMEO on the VIMEO Site and/or in this TOS. Any Submission you provide VIMEO will be deemed public information."* [42]

- *Granting unsecured rights to 3rd party content:* e.g. where this has been collected subject to a licence agreement. For example the JISC *News Film Online* service[43] provides authorised users in HEI's with access to thousands of hours of news clips that may be used for educational purposes, under licence from rights holders ITN (Independent Television News) and Reuters. This was a useful source of experimental stimuli early in the *DIEM* eye movement project, but had to be ruled out as a candidate for publicly accessible archive content since the end-user licence excludes its re-distribution on the open web.
- *Fair dealing* (or 'fair use') exemptions to copyright limitations were an area of concern in the case study to researchers making naturalistic recordings in 'the field', where there is often the possibility of incidentally recording performed music or video. In the UK, unlike the USA, there is no legal right to 'fair use' for (e.g.) non-commercial purposes; it is only a potential defence against litigation. According to the PRIMO repository ('Practice as Research in Music Online')

  > *"The concept of 'fair dealing' is believed to extend to short extracts of copyright printed and recorded music, and film, used expressly for the purposes of criticism and review. This means that you do not need to seek permission to use brief extracts of copyright material (generally accepted as 5% or less of a complete piece, where*

---

42 Vimeo.com Terms of Service at: http://www.vimeo.com/terms (May 12 2009)

43 News Film Online available at: http://www.nfo.ac.uk/ (May 12, 2009)

'complete' relates to a single song, or movement). You cannot invoke the 'criticism and review' exemption if you are using material as illustration. In all cases the original source must be acknowledged, and the extract must not be longer than is necessary for the point to be made."
(http://primo.sas.ac.uk/eprints/cgi/wiki/Copyright_and_ethics 12 May 2009)

- *Performance rights*: where actors or volunteers performing a script or scenario are the subjects of the video data, they may have performance rights unless a waiver is obtained. According to the *Web2Rights* project:

  "*Where an individual does appear in a video clip then she may have performers rights. Performers are not defined in the legislation but performance means 'a dramatic performance (including dance or mime), a musical one, a reading or recitation of a literary work or a performance of a variety act or any similar presentation'. Where the right exists the performer has the right to object to fixation, reproduction, distribution, public performance and communication to the public of the performance without permission. Therefore if the participant can be classed as a performer consent for these activities needs to be obtained.*"[44]

- *Database rights and licensing:* the AMI Project license is similar in terms to the Creative Commons 'Attribution-Noncommercial-Sharealike v2.0' license, but also asserts *Database Rights.* The Database Right is provided automatically to the 'maker' of a database, a term which is not restricted to database management systems and should apply to any corpus:

  "if there has been a substantial investment in obtaining, verifying or presenting the contents of the database (even if the contents and/or structure of the database are not original and therefore do not attract copyright). Investment is construed widely and covers financial, human and technical resources"[45].

- *Open Data licensing* Copyright should be applicable to multimodal research data such as video in which intellectual effort has been invested. However there are jurisdictional differences and different legal views on the extent to which research data may be 'factual' and therefore not copyrightable. Data licensing is therefore an evolving area and the *Guide to Open Data Licensing* [46]is a useful resource for tracking it.

---

[44] Web2Rights Data Protection FAQs at http://www.web2rights.org.uk/navigator/content/documents/4.8%20DP%20FAQs.doc (May 12 2009)

[45] DCC Legal Watch Paper *IPR in Databases* available at: http://www.dcc.ac.uk/resource/legal-watch/ipr-in-databases/ (Aug 5 2009)

[46] Open Knowledge Foundation Guide to Open Data Licensing, available at: http://wiki.okfn.org/OpenDataLicensing

General sources of advice on licence issues affecting video are the BUFVC (British University Film and Video Council), which operates an information service to UK Higher Education Institutions[47], and JISC Legal Information.

3.2.2.1 Ethics and Data Protection

The risks encountered in the study were of obtaining and using personal data without adequately informed consent, both an ethical issue and a legal one where identifiable personal data is disclosed for an unauthorised purpose.

Researchers were well aware of the problematic aspects of video and the additional steps that may be needed to negotiate consent. Appendix 1 gives some examples of these in the FAQ and Consent Form from the *Habitable Cars* project. More general sources of information on access to personal data are listed in box 1.

Sources of advice specific to video data are very limited, especially on limits on recording in public places written consent cannot be practically obtained. The Data Protection Act 1998

Sources of Guidance on Data Access and Confidentiality

Social Research: UK Data Archive at:

http://www.data-archive.ac.uk/sharing/confidential.asp

Clinical Research: UK Clinical Research Collaboration Regulatory & Governance Advice Service at http://www.ukcrc-rgadvice.org

Caldicott Guardian Manual 2006:
http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_062722

NHS Code of Confidentiality:

http://www.dh.gov.uk/en/Policyandguidance/Informationpolicy/Patientconfidentialityandcaldicottguardians/DH_4100550

MRC: Personal Information in Medical Research:

http://www.mrc.ac.uk/Utilities/Documentrecord/index.htm?d=MRC002452

NHS Information Governance: Guidance note on Security of NHS patient data shared for research purposes

https://www.igt.connectingforhealth.nhs.uk/WhatsNewDocuments/NHS%20IG%20-%20security%20for%20research%20guidance.pdf

National Information Governance Board (NIGB) (formerly PIAG) http://www.nigb.nhs.uk/

sets out the general principles, but only the UK Information Commissioner's Office's *CCTV Code of Practice* offers guidelines relevant to video recording in the field. This "covers the use of CCTV and *other systems which capture images of identifiable individuals or information relating to individuals* for any of the following purposes:

- Seeing what an individual is doing, for example monitoring them in a shop or walking down the street.
- Potentially taking some action in relation to an individual, for example handing the images over to the police to investigate a crime.
- Using the images of an individual in some way that will affect their privacy, for example passing images on to a TV company."

---

47 British Universities Film and Video Council available at http://www.bufvc.ac.uk/services/index.html

(UK ICO, 2008, emphasis added)

The Code of Practice applies the DPA to recorded moving images; requiring clearly documented procedures for how the images should be stored and handled, including disclosures. The Code requires that people being recorded are informed, using signs that are an appropriate size depending on context, are clearly visible and readable, and contain details of the organisation responsible, the purpose of the recording and who to contact.

Underlining the need for consent, individuals may also have a right to privacy that extends beyond Data Protection legislation. According to the JISC *Web2Rights* project:

> *"Two areas of law are relevant to the use of images whether in video clips or otherwise. The first is breach of confidence, which has been developing into a right that might be described as a 'right of privacy'; the second is Data Protection law. Broadly, combining these two means that the image of an individual should not be used in a video clip or otherwise unless consent has been given."* [48]

3.2.3 Data Collection & Development – 'Sheer Curation'

The *Data Collection and Development* section of a Data Management Plan broadly identifies what the 'data' will comprise for the research project, describing the data to be collected or captured along with the technologies and methods chosen, the expected volumes and types of data, and quality criteria to be used. It should also identify metadata standards and data formats, any interoperability issues, and methods to be used to integrate the data with any pre-existing sources used.

The uncertainties and challenges identified in Chapter 2 typically stemmed from the difficulties predicting volumes and sources in exploratory projects, the lack of *accepted* metadata and archival format standards, and the high rate of change in capture and storage technologies.

Wider sharing of data can help in mitigating the risks these pose later in the curation lifecycle; one of the benefits attributed to data sessions for example was that they offered an opportunity to pick up the 'craft skills' needed to chose the most appropriate configuration of capture equipment and formats for the setting. The main risks - of unmanageable costs and technology dependencies– should also be mitigated by borrowing from standards for file formats and metadata that fit with research aims, current practices, and the changing environment – and by maintaining current awareness through authoritative sources.

The 'sheer curation' in the subtitle refers to the idea that curation activities should be quietly integrated into the normal workflows of those creating and managing data and other digital assets[49]. This section proposes two main ways to pursue that:

- Building on the proximity of eye movement studies and video ethnography to applied linguistic domains, and adapting their metadata standards and tools.
- Building metadata capture and documentation around the learning aspects of data sessions.

File formats are considered first.

---

[48] FAQ on 'What rights do participants have to their image (likeness)? at www.web2rights.org.uk/navigator/content/documents/4.8%20DP%20FAQs.doc

[49] Sheer curation Wikipedia reference available at: http://en.wikipedia.org/wiki/Digital_curation#Sheer_Curation (12 May 2009)

Formats from Capture to Archiving

The variety of file formats supported is a major cost factor in data archiving according to Beagrie et al, who advise that "dealing with a small number of widely understood file formats allows for simpler procedures at the time of acquisition and future migration" (Beagrie et al 2008).

A degree of duplication is needed though, because compression requirements are different for *archiving*, where the digital master should retain as much as possible of the original, using lossless encoding and resulting in large files, and for *online dissemination* where file sizes need to be minimal especially if the video is to be streamed.

At the *creation* stage, the choice of video capture formats will need to take account later analysis requirements. There is a risk here of experimental results derived from video processing techniques being affected by compression artefacts. For example eye movement researchers in the *DIEM* project intended to use statistics from image analysis of edges and movement, characteristics that are susceptible to artefacts from JPEG2000 and MPEG-2

Best Practices in a Nutshell

- Maintain clear documentation of rights, terms of use, and access restrictions.
- Make an archival copy in a format which offers LOTS (i.e., it is Lossless, Open standard, Transparent, and Supported by multiple vendors).
    - For textual material, use txt format with XML markup, and encode the characters in Unicode.
    - For video, minimize compression using JPEG2000 (J2K) or DV for archival copies.
- Create metadata for the resource in a standard format, using a Dublin Core -based standard. For video use a platform that automatically extracts technical metadata and enable it to be stored as XML text.
- Make the metadata harvestable to a general search engine, even if the resource itself is not available online.
- Plan and budget for storing archival copies in an institutional repository, national data centre, or a stable online community archive that will:
    - Maintain constant unique identifiers for data
    - Migrate data to new formats.

(Adapted from E-MELD Project http://emeld.org/school/bpnutshell.html)

compression (respectively) if lossy compression is used when encoding the files. This was taken into account early in the project when considering storage and equipment needs, since high-definition camcorders may encode in MPEG-2 at the point of capture.

For *archiving* consensus has recently emerged around *losslessly-encoded JPEG2000 wrapped in MXF* as the preferred archival format; it is recommended by the UK Data Archive as the format for depositing ESRC-funded video data. JISC Digital Media also recommend JPEG2000, although they also point out that video-editing software currently lacks support for it, and *DV format wrapped in .mov or .avi* is considered an acceptable alternative. Prestospace recommends migrating DV files to JPEG2000 when the DV format becomes obsolete, as is likely within 5 to 10 years[50].

For *web dissemination* purposes the mp4 format, also widely supported for podcasting, is preferred, encoded using the H.264 codec.

---

[50] PrestoSpace at: http://wiki.prestospace.org/

Authoritative sources on video file formats include *JISC Digital Media*[51] and the *US Library of Congress*[52]. Specific advice on preservation is available from the *Stanford University Video Preservation*[53] page and the EU-funded PrestoSpace[54] project.

Quality and its Storage Implications

Criteria for video data quality will include the image clarity and audio fidelity. While criteria for determining the relevance of the recorded material to unfolding research themes may be difficult to set at the beginning of a project, early decisions on the required clarity and fidelity will pay off in terms of greater control of storage costs, and material that can be shared.

Case study projects used several strategies for limiting storage needs; choosing at the outset to limit file sizes at the capture stage by reducing the video frame dimensions and the audio sampling frequency. It is normally desirable though to retain as much of the original quality as possible. One experienced researcher mentioned the downside of reducing audio quality – this can limit *other* researchers' ability to understand recorded conversation. As well as not having the benefit of being present at the recording, they may lack familiarity with the setting, and may be less familiar with the terminology or local accents and so on.

Wider availability of archived video data and documentation should enable a virtuous circle; by identifying what they need to understand others data, researchers should be better able to identify what will be required to make their own data comprehensible to others *and* better plan resources to meet their own research aims.

Documenting Contextual Information as a Learning Activity

> *All that context… I actually wasn't around for the event, and I didn't video it. I just used other people's video's. What I ended up doing was basically spending a long time chatting to people who were on the project and who were there to build up a proper picture of what was there. And it made me realise how much of a big job it is. And if you are the person who has recorded the video, then I think it's probably even harder for you to actually write all that down. So maybe if you want to store context, it's a case of someone who isn't involved trying to establish the context. That might be a better way of doing it.* (Research Assistant, int. 15)

Researchers using ethnographic approaches in human geography or computing do not have an accepted metadata standard to describe their collected material. Neither do psychology researchers doing experimental eye movement studies. The starting point for the case study groups must lie in existing standards. The questions of how to document contextual metadata in a more standardised way, and using *which* standards, both depend on who the research data is to be shared with (or through which intermediaries) and for what benefit.

The *video ethnographers* for example were sharing informally through data sessions mainly for learning purposes, i.e. to develop their analytic skills and keep track of each other's work. Developing the point in the quote above, the most 'pain free' way to document video that has been shared (or clips of which have been shared) in a data session may be to do it through dialogue between researchers and graduate students, so that students learn by documenting researchers' data.

Currently description is mainly in the form of unstructured notes produced in Word form, relating to the clips or their transcripts, and/or in the (time-coded) logging stored by video editing tools. The most important step to improving data sharing ability – firstly with project

---

51 JISC Digital Media at: http://www.jiscdigitalmedia.ac.uk/movingimages/

52 Library of Congress at: http://www.digitalpreservation.gov/formats/content/video.shtml

53 Stanford University Video Preservation at: http://cool-palimpsest.stanford.edu/bytopic/video/

54 PrestoSpace at: http://wiki.prestospace.org/

collaborators – is to use XML as the storage/exchange format and to use tools supporting relevant XML-based standards. This does not exclude MS Word, for example using a customised document properties in a template.

In principle the MPEG-7 standard is relevant, as it is intended for tagging moving image data with structural and semantic descriptors at a highly granular level, but it is very specific to video data (as opposed to the various other sources of timed data that researchers were working with). Also it is not currently widely supported (LoC, 2009) and video data management platforms that support it are still an active research area (e.g. Küçüktunç, 2008)

The standards used should combine:

- *Institutional repository* or digital library metadata: applied at the level of the corpus or collection, this will be administrative and descriptive metadata based on *Dublin Core*, as used for example in Edinburgh's Datashare repository.
- *Technical audio/video metadata*: collected at the 'item' or file level automatically by video editing/capture software and stored preferably in MXF format. In the case of eye-tracker data this
- *Descriptive metadata*: applied at the 'session' or experiment level, and borrowing from the IMDI standard developed for applied linguistics (see below).
- *Packaging*: METS (Metadata Exchange and Transmission Standard) is the most widely adopted by digital libraries and, if submitting to an institutional repository or data centre this will normally be provided by the repository platform (e.g. Dspace, Fedora) transparently to the research user.

Both the eye movement studies and ethnomethodological video analysis fields are close to linguistics, via studies of reading in the first case and conversation analysis in the second. The *IMDI* (Isles Metadata Initiative) standard, published by the Max Planck Institute, has been used by sociolinguists and ethnographers and provides 'session' descriptors that might also be adapted for describing eye movement experiments.

A *session* in IMDI corresponds to data (audio, video, annotations etc.) relating to a set of actors/ participants in the same location and time, and information about these comprising:

- *Session*: name, title and recording date, location and brief description of the circumstances under which the data was collected.
- *Project*: name, contacts and description;
- *Content*: type/genre, task, modalities, subject, level of participant interaction, degree of planning/scripting, involvement of the researcher, social context (public, private, family, controlled environment), event structure, channel (f2f, human-machine, phone) languages used, description;
- *Actors*: demographic and contact details, role in session, description, languages;
- *Resources*: linked media files, annotations, format information;
- *References*: cross-references to other sessions, related research documents and publications.

An advantage of IMDI is the availability of an editing tool (*IMDI Editor*) that can be tailored to produce custom profiles, and generates XML files. IMDI is used in the *AMI Meetings Corpus*, to enable searching and browsing of structured information about meetings, and the project has developed scripts to generate the metadata automatically from the content and its linguistic annotations (Popescu-Belis and Estrella, 2007). IMDI has also been adapted for the *CLAPI Workbench*[55] at the University of Leon (Bert et al, 2008).

---

[55] CLAPI Workbench: http://clapi.univ-lyon2.fr (18 May 2009)

3.2.4 Access, Data sharing and Reuse

Under this heading a Data Management Plan should describe how data will be made available, the dissemination platform to be used, the process that will be put in place for gaining access, and what permissions and restrictions are to be applied.

The choice of dissemination platform will depend on what finding aids and reuse capabilities are to be offered to 'external' users; i.e. what may be browsed and searched online (content, annotations, metadata) and what may be downloaded. The institution's data repository, digital library, podcasting service, content management system or virtual learning environment may provide audio-visual browsing and searching. Although the capabilities on offer will inevitably be generic and may be oriented to delivery of e-learning or institutional promotion, they may be a valuable channel depending on the project's target audiences and (again) their reuse requirements. For example *DIEM* project eye movement researchers used the University of Edinburgh news pages to show a sample visualisation of viewers watching US President Barak Obama's inauguration ceremony[56] as a means to raise the profile of the project.

Decisions on how data will be shared, accessed and used will take into account common practice in the research community and the limits on access already discussed under ethical and legal issues. Research funders and journal publishers increasingly expect grant holders to have explicit and transparent data sharing policies, as required by the ESRC, including in medical areas where data may be very sensitive. Data sharing plans do not equate to putting all data in the public domain. Video data may be subject to similar limitations as the UK Data Archive routinely applies to the datasets it holds[57], i.e.:

- Providing access to approved researchers only.
- Requiring data access authorisation from the data owner prior to release.
- Placing confidential data under embargo for a given period of time until confidentiality is no longer pertinent.
- Providing secure access to data, limiting the ability to download certain kinds of data or specific sequences.

The limitations imposed will need to be consistent with institutional or departmental data security policies, and the access services provided for the institutions network.

3.2.5 Short-term Storage and Data Management:

The Data Management Plan should identify the anticipated storage volumes, locations and media, together with the responsibilities for managing this and backup storage, and for ensuring that access complies with the data sharing policy. It should also identify how long data needs to be retained for; 'short term' in this context means within the life of a typical (2-5 year) project. The funding body may contractually define retention. Generally, the RCUK Policy and Code of Conduct on the Governance of Good Research Conduct states:

> *"…after the completion of the research: data should normally be preserved and accessible for ten years, but for projects of clinical or major social, environmental or heritage importance, for 20 years or longer" (RCUK, 2009)*

The main risks identified here were of lack of budget control for storage, given the wide range of factors affecting video file size – and the change in technologies that govern many of the factors, especially the availability of high-definition, hard-disk camcorders. As already mentioned, limiting the number of supported file types will help control both technology and

---

[56] 'Study shows viewers have eyes for Obama': http://www.ed.ac.uk/news/all-news/obama-020209 (2 Aug 2009)

[57] UKDA "Manage and Share Data" at http://www.data-archive.ac.uk/sharing/whyshare.asp

staffing costs. Advice can be sought, from the sources mentioned, on the choice of capture technology and source file formats. Online guidelines[58] and file size calculators[59] can help make reasonable estimates based on the amount of footage to be captured, and the required frame size and data rate.

Backup storage

Presenting data at conferences and data sessions requires removable storage, and the popularity of cheap external drives reflected this in the study. Removable drives have hidden costs however; their reliability is uncertain (and anecdotally low) and time taken to manage backup and recovery might be better spent if there is a departmental backup service. Overall costs of the archival (LTO or 'Linear Tape Open') tape normally used by backup services are lower than for disk storage, although backup tape is much slower to access. Where hard disk-based camcorders are used, LTO tape may be a viable alternative to the miniDV tapes that researchers in the case study projects have typically used as an archival backup. Using writeable CDs and DVDs, on the other hand, is not recommended for archival purposes.

The estimates for storage should take account of archival best practice, which is to store at least two 'master' copies of each media file, and in separate locations. This doubles the overall capacity needed for archival copies, and web (dissemination) copies will add further.

3.2.6 Deposit and Long-term Preservation

Institutional and national data repositories are moving to deal with sound and moving images, however even specialist archives experienced in dealing with video data, e.g. in the language documentation field, acknowledge the difficulties it presents;

> *"Creating, processing, and storing video is vastly more costly in every imaginable way compared to other media (except perhaps interactive multimedia). It is more intrusive in the field situation … [and] it is ravenously hungry for resources (human and machine) in editing, production, distribution, and storage, because the technologies for digital video are complex, in constant flux, and data volumes are extremely high*. (Nathan, 2008, p.3*)"*

Progress in supporting research needs will therefore depend on partnerships between data support services and researchers in video analysis who are experienced in dealing with these issues. Researchers meanwhile are expected to identify a *long-term preservation and curation strategy* in their Data Management Plan. This will cover steps to be taken towards the end of the project; the data transformations needed for long-term reuse.

The long-term strategy will cover similar activities as those already identified for 'project curation': identifying the retention timescale, how data will be selected or disposed of, the anonymisation procedures and access/sharing limitations to be enforced, the additional metadata that will be added, steps that will be taken to migrate the data to avoid technology obsolescence, and where and how data will be deposited in a public repository. Offering data to an archive that caters specifically for the research community may also be desirable, and may be the only option available if the funding body has no mandated data centre.

Institutions are likely to develop services that perform specific curation and preservation functions to defined service levels and costs, as they currently do for e-learning. One example these may follow is the 'self-archiving' facility piloted by UK Data Archive, *UKDA Store*. This is currently restricted to ESRC-funded researchers, offering them the capability to upload "a range of digital objects to the repository including statistical data, databases, word documents, audio-visual materials", and to assign access permissions to individuals and/or groups. Users can also ask for uploaded data to be 'formally lodged' with UKDA, i.e. self-

---

[58] Adobe.com "Calculate File Sizes and Disk Space for Video" at: http://kb2.adobe.com/cps/312/312640.html

[59] e.g. VideoSpace for Mac OSX at: http://www.digital-heaven.co.uk/videospace

archiving does not replace the traditional deposit route or the added-value services (such as quality assurance and migration) offered for accepted datasets.

Owners of multimodal research data collections or corpora may want to check that the archive they will be using supports the *metadata requirements for video preservation*. Metadata based on the Dublin Core (DC) standard are normally created by an archive at the point of ingest. These record descriptive and contextual (administrative) information about the 'collection' or corpus such as its title and creator. DC metadata is necessary but may not be sufficient for video preservation. To retain the 'significant properties' other information specified by the PREMIS and MPEG-7 standards is needed; according to the *recommended minimum element set* identified in the *Moving Images and Sound Archiving Study* for the Arts and Humanities Data Service (Wilson et al 2006). Much of this is technical metadata, such as the aspect ratio and sampling format, which should be retrievable from the video wrapper file, preferably in XML form. Others are normally created by an archive- such as the 'checksum' codes added to files to maintain their '*fixity*' or authenticity, and the unique identifiers used to link all the files that comprise a particular version of the collection or corpus.

Video format migration – *transcoding* – is also a service that might be provided for large video collections. Even although desktop software such as VLC have transcoding capabilities this may stretch computational resource - any archive wishing to migrate hundreds of hours of losslessly-compressed master copies is likely to need support, at least in computation services if not organisationally.

Large-scale storage that is offered as a centrally managed service will likely demand some form of quality assurance. The case study suggests that the main criteria for reviewing or appraising video and other multimodal data in social research – apart from on quantifiable attributes such as clarity and presence or absence of metadata and documentation - is likely to be their relevance to the research theme rather than anything more 'objective'. This criterion may only be applied towards the end of projects, to the data that has been selected for analysis. This implies a storage hierarchy, topped by data directly related to a research publication, followed by anything else selected analysed, with the remainder of captured data (which as we have seen may be 90% of the total) remaining the responsibility of the original research group and having a limited lifespan post-project.

The 'community curation' scenario, where approved researchers can browse and comment on video their peers have annotated and made available, is more likely to be sustained over the short term – when dialogue with the original researchers is feasible – but should also be a feature of long-term curation by institutional data repositories. If it is to be sustained, a model such as the *VidArch* project's Contextual Information Framework or the *Diver* project's 'webdiver' will be needed to underpin it.

3.2.6 Curation Phases and Researcher Roles

The chapter began with a phased approach to the DCC Curation Lifecycle, illustrated in Figure 3.1. Roles and responsibilities for these phases have not been identified except that implicitly a progressively greater role would be played by research data services as these are developed institutionally and nationally. It was not in the scope of the case study to attempt to define roles and responsibilities, but Table 3.2 below describes for the purposes of further discussion how these might evolve, involving three main groups of actors:

- *PI & Research Group* i.e. academics and researchers directly involved in creating or capturing research data.
- *Computing Support* at a devolved (e.g. School or Department) level of the institution;
- *Research Data Services* provided through the Institution, the funding body (e.g. ESRC) or UK-wide organisations such as DCC, JISC Digital Media or JISC Legal. Institutional services for supporting audio-visual media are likely to offer expertise and advice on copyright, IPR and technical choices, although services are likely to be oriented to producing e-learning materials rather than research data.

Planning and Piloting First iteration of curation lifecycle: from project bid stage through to the researchers' organisation of acquired data, with limited data sharing, backup managed at School/Dept level, support from Research Data services in preparing Data Management Plans, risk assessment, and advice on metadata and standards, tools and curation infrastructure. **Roles**: *authorised* by PI & Research Group, *performed* by PI & Research Group and Computing Support, *advised* by Computing Support and Research Data Services. **Characteristics**: Data plans established, main risks identified and assessed, video data captured and backed up in original or 'digital master' format, other time-based data collected, some selected and encoded, content logged, rough transcripts, notes documented.	**Formats/ standards** DV + .mov or .avi JPEG2000 + MXF TXT, RTF **Tools** Video editing & logging, word processing or blog. Experimental workflow modelling
Project Curation Second iteration: from re-assessment of scope of data as first conceived, and risks to its re-usability; develop the Data Management Plan, carry out the bulk of capture, annotation and analysis. Involves more collaboration with colleagues or peers, and preparation of extracts for wider publication and archiving. Uses a collaborative data store managed by the Research Group or at School/Dept level, supported by Research Data Services. **Roles**: *authorised* by PI & Research Group and Computing Support, *performed* by PI & Research Group, Computing Support, Research Data Services, *advised* by Computing Support, and Research Data Services **Characteristics**: Data management procedures established, colleagues/ peers and (possibly) research participants involved in annotating and assessing data, more descriptive and context information recorded, links to transcripts & associated data established, selected data analysed further, transcoded to distribution formats	**Formats/ standards** As above, plus IMDI – session metadata (descriptive, administrative) MPEG4, H.261 (web dissemination) **Tools** Annotation e.g. DRS, NXT, ELAN, Diver/ Webdiver Visualisation (e.g. eye tracking) Metadata editor e.g. IMDI Content/ Asset Mgmt System or repository (e.g. Dspace, ePrints, Fedora)
Long-term Curation and Preservation Further iteration(s) of the Curation Lifecycle, as the PI & Research Group and Research Data Services appraise and select data and negotiate deposit/ingest or transfer to a public or community repository within the Institution or elsewhere, and specify preservation, access and reuse requirements. **Roles**: *authorised* by PI & Research Group, and Research Data Services *performed* by PI & Research Group, Computing Support, Research Data Services, *advised* by Computing Support, and Research Data Services. **Characteristics**: Risks re-assessed, data quality re-appraised, metadata augmented through ingest process, persistent identifiers defined, access limitations defined, selected data transcoded as required, data and annotations made accessible and linked to publications, community annotations and reuse information collected and moderated according to specifications.	**Formats/ standards** As above, plus Dublin Core (descriptive, admin), METS (structural) **Tools** Transcoding Content/ Asset Mgmt System or repository (e.g. Dspace, ePrints, Fedora) Web video search/browsing and annotation (e.g. Webdiver)

Table 3.2 Curation phases and possible roles

4 Conclusions and Recommendations

The case study reviewed the curation landscape in several interdisciplinary fields that use video analysis in studies of human interaction. Field research for the study involved semi-structured interviews with researchers and participant observation in their informal data sharing and analysis sessions. Researchers were working in a variety of projects across social sciences, humanities and computing; from domains including human geography, psychology, sociology, social informatics, and linguistics. The study primarily focused on uses of video in ethnographic studies and in eye movement research.

4.1 Video in Interaction Analysis

Research projects in interaction analysis tend to be funded by a variety of bodies, and organised either around the lone researcher, or through interdisciplinary projects that may span departmental and institutional boundaries. Researchers use informal networking with peers for support with analysis techniques and technology issues, often relying on department-level support for managed storage and web publication, with less use of institutional-level services where support for video data archiving is embryonic.

The projects featured in this study were generally aiming to improve understanding of human interaction, with potential application in (for example) technology design and language learning. Some of the projects featured were building or using tools enabling observational data of different communication modalities to be synchronised and analysed together. In eye movement research, data is being visualised and made available to other research communities to stimulate the development of modelling techniques. Tools that synchronise timed data and visualise the results help to facilitate multi-disciplinary collaboration. This in turn drives curation - in terms of the need for shared storage, access and terminology.

Research material in these fields consists of collections (corpora) of multimodal data i.e. video plus audio, synchronised with a range of other interactions (e.g. eye movements of research participants watching video clips), and/or derived data (transcripts, annotations, analyses). Corpora were being used as observational records for ethnographic analysis, and in eye movement research they comprised the video used as experimental stimuli, together with the resulting eye tracking data and video visualising these. Researchers were using video captured from camcorders in DV or HDV format, or collected from online sources, encoded according to various MPEG-4 compression standards and stored in QuickTime or AVI wrapper formats. Related data such as eye movements are typically recorded with time codes to enable synchronisation, and stored in text files. There are few community initiatives to standardise these.

The typical research cycle, after the design stage, involves firstly finding and recruiting subjects or settings. In ethnographic studies the ‘subjects’ will to varying degrees be active participants in research and will in any case influence the selection of video clips at the point of capture. The research is ‘data driven’ in that data and research themes co-evolve in the course of a project. The more experimental domain of eye movement research is also data driven, at least in studies of the perception of moving images, since there is a lack of models to hypothesise from. In both forms of research, early stages of analysis further select and reduce the data by identifying patterns, annotating, linking and re-organising the data according to their relevance to emerging themes. Visualisation of interaction patterns in and between video and other data provides a foundation for hypothesis testing in experimental studies and richer description in ethnographic analyses.

At the creation stage, consumer oriented video editing tools were used for organising and logging selected clips. Researchers using ethnomethodology and conversation analysis approaches used common word processing tools for transcription, although more specialised tools are available for this. Qualitative data analysis tools have generally only recently offered

capabilities to annotate video data at sufficient levels of granularity for interaction analysis, where more relevant tools have been transferred from linguistic domains. Several of the projects featured in the case study were developing tools with sophisticated coding and annotation capabilities for multimodal data, and to visualisation patterns for eye movement studies. There is a lack of tools that assist researchers to manage video data across its lifecycle and through transitions to web collaboration and longer-term curation.

Much informal data sharing is carried out without structured or standardised metadata. Video data is frequently shared informally with known peers, for example in 'data sessions'. Researchers benefit from colleagues' contributions to the analysis, and also from the learning and development opportunities data sessions afford. Analysis in these sessions (and by implication subsequently) is shaped by practical data storage limitations (i.e. what can easily be transported) and time limits (i.e. what can easily be presented for comment). The data description that accompanies sharing at this stage is (necessarily) unstructured, but is a source of contextual information.

Technical metadata that allow encoded video and audio data to be rendered tend to be captured automatically by the originating software, and are broadly consistent with defined 'significant properties' for preservation. While standards exist for descriptive metadata (e.g. MPEG-7) they are not widely supported in current tools for video data management. Descriptions, or derived data, instead use a variety of community-specific schema, ranging from notation and markup inserted into common word processor formats, to XML-based coding schema for linguistic annotation. Annotation (coding) schemas tend to be project-specific where interpretive analysis is used, and in some fields they are avoided. Formalised annotation and metadata schemas are more common in sociolinguistic domains, and may have wider take-up through interdisciplinary collaborations.

4.2 Enablers and Constraints on Discovery, Access and Use

Web-based video is universally available and in great quantities. Easy-to-use hosted web video platforms are often used for research communication, although rarely for data. There are very few community archives or repositories that provide the capabilities to browse and search online video and related material on social interaction, and the exemplars are mainly in the US and elsewhere in Europe.

Access limitations are legal, organisational and ethical as well as technical. Confidentiality is a major issue since anonymising video is far more challenging and costly than is text. The risks that public access poses to subjects are regarded as high, and consent may be impractical to obtain. IPR in video data also affects its potential for reuse. At point of capture, there is lack of clarity over what is permitted under 'fair dealing'. Resources licensed for educational use that are used for research are not openly re-publishable as research outputs (e.g. on a 'copyleft' basis). Researcher using commercial 3rd party web video platforms risk transferring to them IPR they have not secured.

Technical and organisational issues relate to *search*, since video content is only searchable to the extent it has been transcribed or annotated in other ways; and to storage and bandwidth; since detailed interaction analysis requires high-resolution video, which is costly to manage and to transfer. Storage requirements are affected by changes in media used at point of capture, i.e. disk-based camcorders, and the need to maintain high quality digital masters. Short-term storage requirements may be far higher than long-term, since most collected material will be discarded as irrelevant.

Enablers include the availability of relevant metadata schema. European linguistic research groups such as the Max Plank Institute for Psycholinguistics and the ICAR Research Lab (CNRS and University of Lyon) provide metadata schema (e.g. *IMDI*) that may be transferable to other forms of interaction research. The linguistic community has well-established multimodal archives, which are developing frameworks for online browsing and annotation that would provide a step-change in the reusability of video data, along lines

proposed in US research projects such as *VidArch* and *Diver*.

Interaction analysis of multiple synchronised streams of data needs sophisticated tools to allow multiple layers of data and coding to be visualised and searched. UK research groups including those featured in the case study are leading development in this area. Examples here include *Digital Replay System* (DRS), the *NITE XML Toolkit* (NXT) and visualisation software produced by the Edinburgh Eye Movement Studies group.

Institutional or national data repositories are only beginning to publish online information on the availability of video data, and rarely enable searching across video-specific metadata or annotations. Enabling work in this area includes the UKDA *DEXT* project which, though not specific to video data, provides a generalised XML scheme for qualitative data annotation, with a level of automated annotation. This might be extended to support automatic metadata extraction from annotations, following the example of the *AMI Meeting Corpus*.

Recommendation 1- DCC, JISC and other research funders should develop the e-infrastructure for multi-disciplinary interaction research by facilitating workshops to bring together the disciplines involved, disseminate relevant tools, and explore more effective ways to browse and annotate multimodal data in data repositories. This would take forward work piloted by UKDA and by the DreSS project.

4.3 Data Management Planning and the Curation Lifecycle

Projects featured in the case study had carried out data management planning to funding-body requirements. Researchers considered the DCC Curation Lifecycle Model helpful to this, however data management plans for projects that are dependent on video data can quickly get out of step with reality. Video is voluminous, capture technologies and formats are changing, and there are many interdependencies between storage management and other stages in the curation lifecycle. The lifecycle may be more helpful if applied iteratively.

The case study provides a phased approach to data management planning using the Curation Lifecyle model iteratively. This involves re-assessing the curation risks, at the points where changes in curation activity align with changes between the phases of a project; the early ‘planning and piloting’ stage where data is typically managed according to individual researchers’ need, the main body of the project when it may be re-organised to support collaboration, and the later stages of publication and dissemination when (post-project) long-term curation needs are of greater concern.

Recommendation 2 –The DCC Curation Lifecycle Model is an ‘ideal type’ and rather than used as a one-off framework for Data Management Planning it should be used iteratively during research projects, by periodically reviewing the Data Management Plan so that research materials that have been collected can be used effectively by the core research team, collaborators and other potential reusers.

Researchers initial selection and appraisal of data was informal; exploratory research methods were mostly used in the projects studied; provided that video had enough clarity and audibility to enable interpretation and analysis at the desired level of granularity, it’s quality as research evidence would depend on the themes or hypotheses to emerge from early analysis. For observational data, this depends on the provenance, precision and interpretability of the data and annotations. Assessing this involves at least replaying the data to follow a path from its observable features to the researchers transcription and interpretation, to assess the detail and plausibility of the analysis and its relevance to themes of interest to the research community. In some domains frequency counts may be made of annotations as a check on reliability and validity. Data quality assessment tends not to be formalised in interpretive social science research, except in that metadata completeness checks are standardised at the archival ingest stage.

In principle UKDA is the most relevant place for researchers to deposit multimodal corpora. However projects were often interdisciplinary, and had other funders including EPSRC and

AHRC and charitable foundations with no equivalent data centres. Researchers expected more of their institutions in provision of affordable managed storage, technical support on formats, and a post-project 'parking place' for datasets with potentially unrealised value.

Recommendation 3 – DCC should collaborate with relevant Research Councils, JISC Digital Media and JISC Legal Information to guide institutions, research ethics committees, and researchers on planning and managing the curation of video and multimedia research data.

Recommendation 4 – HEI's should consult researchers on the methodological and technical issues affecting the reusability of video and multimodal data they would want to submit to institutional or subject data repositories, and coordinate the support they provide with the relevant services provided by JISC and other agencies.

Acknowledgements

Thank you to all who contributed to the case study, and particularly to Dr Eric Laurier, School of Geosciences, University of Edinburgh for his patience and generosity.

References

Berck, P. and A. Russel. A. (2006). 'Annex - a web-based framework for exploiting annotated media resources'. In *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006).*

Bert, M. Bruxelles, S. Etienne, C. Mondada, L. and Traverso, V. 'Tool-assisted analysis of interactional corpora: voil`a in the CLAPI database' *French Language Studies* 18 (2008), pp. 121–145.

Bertenthal, B., Grossman, R., Hanley, D., Hereld, M., Kenny, S., Levow, G., Papka, M. E., Porges, S., Rajavenkateshwaran, K., Stevens, R., Uram, T., and Wu, W. (2007) "Social Informatics Data Grid,", *Third International Conference on e-Social Science* October 7-9, 2007, Ann Arbor, Michigan

Bishop, L. (2006) 'A proposal for archiving context for secondary analysis' *Methodological Innovations Online* 1(2)

Brown, B. and Laurier, E. (2005) Maps and Journeys: an Ethnomethodological Investigation,. *Cartographica* 4 (3), pp. 17-33.

Carletta, J Hill, R Nicol, C. taylor, T

Carter, R. and Adolphs, S. (2008) Linking the verbal and visual: new directions for corpus linguistics *Language and Computers*, 64, pp.275-291

Corti, L. (2007) 'Re-using archived qualitative data – where, how, why?' *Archival Science* (7) pp. 37–54

Crabtree, A., French, A., Greenhalgh, C. Benford, S., Chevherst, K., Fitton, D., Rouncefield, M. and Graham, C. (2006) 'Developing digital records: early experiences of record and replay' *Computer Supported Cooperative Work* 15 (4), pp. 281-319

Dourish, P. (2004) 'What we talk about when we talk about context' *Personal Ubiquitous Computing*, 8 pp. 19–30

EMCA News (2007) 'Developing Data Sessions' *Newsletter of the American Sociological Association Section on Ethnomethodology and Conversation Analysis* 2(1) Summer 2007

Fielding, N. (2003) *Qualitative Research and E-Social Science: Appraising the Potential* ESRC Consultative Study, Engineering and Social Research Council, UK.

Fraser, M. Biegel, G. Best, K. Hindmarsh, J. Heath, Greenhalgh, C. and Reeves, S. (2005) 'Distributing Data Sessions: Supporting remote collaboration with video data' *First International Conference on e-Social Science*. June 2005

Fraser, M. Hindmarsh, J. Best, K. Heath, C. Biegel, G. Greenhalgh, C. and Reeves, S. (2006) 'Remote Collaboration Over Video Data: Towards Real-Time e-Social Science' *Computer Supported Cooperative Work (CSCW)* 15(4) pp. 257-279.

Friese, S. (2004) "Computer-aided Qualitative Data Analysis of Multimedia Data: Technological Advances, Challenges and Methodological Implications" CAQDAS Networking Project presentation 24th March 2004, University of Surrey available at: http://caqdas.soc.surrey.ac.uk/seminarslides.html

Fry, J. (2006) 'Scholarly research and information practices: a domain analytic approach'

Information Processing and Management 42 (2006) pp. 299-316

Giddens, A.(1984) *The Constitution of Society. Outline of the Theory of Structuration*. Cambridge : Polity

Greenberg, S. (2001) 'Context as a dynamic construct' *Human Computer Interaction* 16 (2) pp. 257-268

Henderson, J.M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, 7, 498–504.

Henderson, J. 'Regarding Scenes' (2007) *Current Directions in Psychological Science* 16(4), pp. 219-222.

Hindmarsh, J. (2008) 'Distributed Video Analysis in Social Research', in Blank, G., Fielding, N. and Lee, R. (eds) *Handbook of Online Research Methods*, London: Sage

Hindmarsh, J. and C. Heath (2007) 'Video based Studies of Work Practice' *Sociology Compass*, 1(1): 156-173.

Holland, J. Thomson, R. and Henderson, S. (2006) *Qualitative Longitudinal Research: A Discussion Paper Families & Social Capital* ESRC Research Group London: South Bank University

Hughes, B. Schmidt, D. and Smith, A. (2006). Towards Interoperable Secondary Annotations in the E-Social Science Domain. *Proceedings of 2nd International Conference on E-Social Science.* National Centre for E-Social Science: Manchester.

Jefferson, G. (1984) 'Transcription Notation' in: Atkinson, J. and Heritage, J. (eds), *Structures of Social Interaction*, Cambridge: Cambridge University Press.

Jirotka, M. and Luff, P. (2006) 'Supporting Requirements with Video-Based Analysis'. *IEEE Software*. 23(3)

Knoblauch, H. (2006) 'Videography: Focused Ethnography and Video Analysis', in: Knoblauch, H., Raab, J., Soeffner, H.-G., Schnettler, B. (eds). *Video Analysis*, Oxford: Lang

Laurier, E. and Philo, C. (2006) 'Natural Problems of Naturalistic Video Data', in: Knoblauch, H., Raab, J., Soeffner, H.-G., Schnettler, B. (eds). *Video Analysis*, Oxford: Lang

Lee, C. (2007) *Taking Context Seriously: A Framework for Contextual Information in Digital Collections*, Technical Report UNC SILS TR-2007-04, University of North Carolina

Lomax, H., and Casey, N. (1998). Recording social life: Reflexivity and video methodology. *Sociological Research Online*, 3(2)

Luff, P. and Heath, C. (2002) Broadcast Talk: Technologically Mediated Action in a Complex Setting, *Research on Language and Social Interaction.,* 35 (3), 232-264.

Lyon, L. (2007) *Dealing with Data: Roles, Rights, Responsibilities and Relationships* UKOLN, University of Bath available at: http://www.ukoln.ac.uk/ukoln/staff/e.j.lyon/publications.html#2007-06-19

Macbeth, D. (2001) On "*reflexivity*" in qualitative research: two readings and a third. *Qualitative Inquiry*, 7, 35-68

McConvell, P. (2003) Multilingual, multiperson, multimedia: linking audio-visual to text material in language documentation. Available at: http://hdl.handle.net/2123/1429 (21 March, 2009)

Marsden, A., Nock, H., Mackenzie, A., Coleman, J., Lindsday, A. and Kochanski, G. (2007) 'ICT Tools for Searching, Annotation and Analysis of Speech, Music, Film and Video: Prospects for Research in the Arts and Humanities', *Journal of Literary and Linguistic Computing* 22: 469-488

Mondada, L. (2006) 'Video Recording as the Reflexive Preservation and Configuration of

Phenomenal Features for Analysis', in Knoblauch, H., Raab, J., Soeffner, H.-G., Schnettler, B. (eds). *Video Analysis*, Oxford : Lang.

Mondada, L. (2007) Operating together through videoconference: Members' procedures for accomplishing a common space of action. In : Hester, S., Francis, D. (eds). *Orders of Ordinary Action* . Aldershot : Ashgate, 51-67.

Nathan, D. (2007) 'Digital Video in Documentation and Archiving' *Language Archives Newsletter*, No.9 Feb 2007 pp. 3-4

Parry, O. and Mauthner, N. (2005) 'Back to Basics: Who Re-uses Qualitative Data and Why? *Sociology* (39) p 337

Pea, R. D. (2006).Video-as-data and digital video manipulation techniques for transforming learning sciences research, education and other cultural practices. In J. Weiss, J. Nolan & P. Trifonas (Eds.), *International Handbook of Virtual Learning Environments* (pp. 1321-1393). Dordrecht: Kluwer Academic Publishing.

Pink, S. (2007) *Doing Visual Ethnography: images, media and representation in research* 2nd edition London: Sage

Popescu-Belis, A. and Estrella, P. (2007) 'Generating Usable Formats for Metadata and Annotations in a Large Meeting Corpus' *Proceedings of the ACL 2007 Demo and Poster Sessions*, pp. 93–96, Prague, June 2007

RCUK (2009) *RCUK Policy and Code of Conduct on the Governance of Good Research Conduct* (DRAFT) Research Councils UK, March 2009, available at: http://www.rcuk.ac.uk/review/grc/default.htm

Randall, D., Harper, R. and Rouncefield, M. (2007) *Fieldwork for Design*, London: Springer-Verlag

Rouncefield, M, Hartswood, M, Procter, R, and Clarke, K (2003) Trusting the record. *Methods of Information in Medicine*, 42 ( ). pp. 345-352

Schroeter, R. Hunter, J. Guerin, J. Khan, I. and Henderson, M. (2006) 'A Synchronous Multimedia Annotation System for Secure Collaboratories' *2nd IEEE International Conference on E-Science and Grid Computing (eScience 2006).* Amsterdam, Netherlands. December 2006. p 41.

Shotton, D. Torbica, D. and Pybus, J. (2003) 'VideoWorks –Blueprint for the functioning of a National Video DataGrid' *Proceedings of UK e-Science All Hands Meeting 2003* 2-4th September, Nottingham, UK

Shotton, D. Zhao, J. and Klyne, G. (2008) *Defining Image Access Project Final Report*, available at: http://imageweb.zoo.ox.ac.uk/pub/2007/DefiningImageAccess/FinalReport/ (21 May, 2009)

Stewart, J. and Hyysalo, S. (2008) 'Intermediaries, Users and Social Learning in Technological Innovation' *International Journal of Innovation Management* 12 (3) pp. 295–325.

Suchman, L. and Trigg, R. (1991). Understanding Practice: Video as a Medium for Reflection and Design. In J. Greenbaum & M. Kyng (Eds.), *Design at Work: Cooperative Design of Computer Systems*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 65-89

Tarrant, D. and Hitchcock, S. (2008) Preservation as a Process of a Repository. In: *The Sun Preservation and Archiving Special Interest Group (PASIG) Fall Meeting*, 18 - 21 November 2008, Baltimore.

Tibbo, H., Lee, C., Marchionini, M., Howard, D. (2006). 'VidArch: Preserving Meaning of Digital Video over Time through Creating and Capture of Contextual Documentation'. *Proceedings IS&T Archiving 2006*

Twidale, M, Rodden, T and Sommerville, I (1993). 'The designers' notepad: supporting and

understanding cooperative design' *Proceedings of the third European Conference on Computer-Supported Cooperative Work*, Milan, Italy. pp. 93 – 108

Van Gompel, R., Fischer, M., Murray, W., & Hill, R. (2007)). Eye-movement research: An overview of current and past developments. In R.P.G. van Gompel, M.H. Fischer, W.S. Murray, & R.L. Hill (Eds.). *Eye movements: A window on mind and brain*. Oxford: Elsevier.

van der Linde, I, Rajashekar, U, Bovik, A, and Cormack, L. (In Press). "DOVES: A Database of Visual Eye Movements", Spatial Vision, 2008. Available at: http://live.ece.utexas.edu/research/doves. (5 Jan 2009)

vom Lehn, D. and Heath, C. (2006) 'Discovering Exhibits: Video-based Studies of Interaction in Museums and Science Centres', in: Knoblauch, H., Raab, J., Soeffner, H.-G., Schnettler, B. (eds). *Video Analysis*, Oxford : Lang.

Wactlar, H., Christel, M., (2002) 'Digital Video Archives: Managing through Metadata'. In *Building a National Strategy for Digital Preservation: Issues in Digital Media Archiving*, Commissioned for and sponsored by the National Digital Information Infrastructure and Preservation Program, Library of Congress.

Wade, N. (2007) 'Scanning the Seen: Vision and the Origins of Eye-Movement Research' In R.P.G. van Gompel, M.H. Fischer, W.S. Murray, & R.L. Hill (Eds.). *Eye movements: A window on mind and brain*. Oxford: Elsevier.

Wesch, M. (2008) 'An anthropological introduction to YouTube' available at: http://www.youtube.com/watch?v=TPAO-lZ4_hU&feature=fvw (12 November, 2008)

Wilson, A., Wright, R., Polfreman, M., Anderson, S., Tanner, S. and Beer, E. (2006) *Digital Moving Images and Sound Archiving Study* London: Arts and Humanities Data Service.

Wittenburg, P., Broeder, D. Kemps-Snijders, M. Dimitriadis, A. and Soddemann Th A (2007) 'Federation of Language Archives Enabling Future eHumanities Scenarios German eScience' Available online at http://www.ges2007.de (21 March 2009)

Wittenburg, P., and Broeder, D. (2008) 'On the Relevance, Standards and Usage of Metadata for electronic language resources' in: Witt, A. Sasaki, F. Teich, E. Calzolari, N. and Wittenburg, P. (eds.) *Proceedings LREC 2008 Workshop on uses and usage of language resource-related standards*.

Zimmermann, A., Lorenz, A., Oppermann, R. (2007): 'An Operational Definition of Context' *Proceedings 6th International and Interdisciplinary Conference (CONTEXT 2007), Lecture Notes in Computer Science 4635*, Berlin: Springer, pp. 558-571

Appendices

Appendix 1. DCC Curation Lifecycle Steps

Full Lifecycle Actions

DATA Data, any information in binary digital form, is at the centre of the Curation Lifecycle. This includes:
Digital Objects: simple digital objects (discrete digital items such as text files, image files or sound files, along with their related identifiers and metadata) or complex digital objects (discrete digital objects made by combining a number of other digital objects, such as websites)
Databases: structured collections of records or data stored in a computer system

DESCRIPTION AND REPRESENTATION INFORMATION Assign administrative, descriptive, technical, structural and preservation metadata, using appropriate standards, to ensure adequate description and control over the long-term. Collect and assign representation information required to understand and render both the digital material and the associated metadata.

COMMUNITY WATCH AND PARTICIPATION Maintain a watch on appropriate community activities, and participate in the development of shared standards, tools and suitable software.

CURATE AND PRESERVE AND PRESERVATION PLANNING Be aware of, and undertake management and administrative actions planned to promote curation and preservation throughout the curation lifecycle.

Plan for preservation throughout the curation lifecycle of digital material. This would include plans for management and administration of all curation lifecycle actions.

Sequential Actions

CONCEPTUALISE Conceive and plan creation of data, including capture method, storage options.

CREATE OR RECEIVE Create data including administrative, descriptive, structural and technical metadata. Preservation metadata may also be added at the time of creation. Receive data, in accordance with documented collecting policies, from data creators, other archives, repositories or data centres, and if required assign appropriate metadata.

APPRAISE AND SELECT Evaluate data and select for long-term curation and preservation. Adhere to documented guidance, policies or legal requirements.

INGEST AND PRESERVATION ACTION Transfer data to an archive, repository, data centre or other custodian. Adhere to documented guidance, policies or legal requirements.

Undertake actions to ensure long-term preservation and retention of the authoritative nature of data. Preservation actions should ensure that data remains authentic, reliable and usable while maintaining its integrity. Actions include data cleaning, validation, assigning preservation metadata, assigning representation information and ensuring acceptable data structures or file formats.

STORE Store the data in a secure manner adhering to relevant standards.

ACCESS, USE AND REUSE Ensure that data is accessible to both designated users and reusers, on a day-to-day basis. This may be in the form of publicly available published information. Robust access controls and authentication procedures may be applicable.

TRANSFORM Create new data from the original, for exampleby migration into a different format, or by creating a subset, by selection or query, to create newly derived results, perhaps for publication.

Occasional Actions

DISPOSE Dispose of data, which has not been selected for long-term curation and preservation in accordance with documented policies, guidance or legal requirements. Typically data may be transferred to another archive, repository, data centre or other custodian. In some instances data is destroyed. The data's nature may, for legal reasons, necessitate secure destruction.

REAPPRAISE Return data which fails validation procedures for further appraisal and re-selection.

MIGRATE Migrate data to a different format. This may be done to accord with the storage environment or to ensure the data's immunity from hardware or software obsolescence.

Appendix 2. Information Sheet & Consent Form for Video Ethnography

(Courtesy of Dr Eric Laurier, School of Geosciences, University of Edinburgh)

Habitable Cars: the organisation of collective private transport

(Info sheet)

Why is this study being done?

Geographers and other researchers have an ongoing interest in how people interact while on the move during short trips or longer journeys. Cars, the particular form of transport we are studying, have an important place in our everyday lives. They are the means for traveling with friends, family and workmates, and transporting shopping, equipment and our possessions. Our aim in this study is to document, describe and report on how people organise traveling *together* in the same vehicle or *sharing* the same vehicle *sequentially*. We are interested in how people interact with one another in cars, what they do whilst in cars and how the car fits into activities that happen outside the car. To do this we will be recording interviews and filming naturally-occurring social activity.

Who is responsible for the data collected in this study?

Eric Laurier is an experienced social researcher who uses ethnographic methods to study various aspects of social life. In the past he was worked on several different topics: urban development, improving health, car-based office work, the design of new technologies, cafes, and, community in suburban neighbourhoods. In this project he is collaborating with Dr Barry Brown (at the Department of Computing Science, University of Glasgow) and Hayden Lorimer (at the Department of Geography, University of Glasgow).

Contact details: *Dr Eric Laurier, Senior Research Fellow,*
Institute of Geography,University of Edinburgh,
Edinburgh EH8 9XP, 0131 651 4303

What is involved in this kind of study?

The study proceeds in two parts over two weeks:

1. **Week 1**, Eric Laurier will travel with you during half a dozen or so of your typical weekly car journeys (once you are at your destination he will leave you and make his own way back). During these journeys, at times when it will not disrupt your travel, Eric will ask a few questions prompted by the journey about these journeys and what occurs during them. These will also be good opportunities for you to learn more about the project. This method of traveling with drivers has been used successfully to elicit information for studies of driver and passenger interaction in the USA.

2. **Week 2**, Using two miniDV cameras secured in the back and front of your vehicle, we will ask you to tape half a dozen of your typical weekly car journeys. We will try to avoid disrupting your everyday travel activities during filming, in fact, our hope is that you will

behave as naturally as possible. Eric will meet up with you daily to reload the cameras with one hour tapes and recharge their batteries. You will have complete control over starting and stopping the recording. After the film has been edited we will contact you, provide you with a DVD, CD or VHS copy of your footage and check that you are happy with the sections of film we would like to use for analysis and presentation. In some cases we may also ask you to help us make sense of the activity captured on the tape.

Who will look at the resulting interview data?

Researchers using these data are members of a professional community that requires them to respect the views of all participants. Segments of the tape with accompanying transcriptions may be presented in the context of the project's results and final report to its funding body[60], scholarly publications, academic symposia, university classes, professional training activities or dissemination of results to the media, policymakers, government and industry.

How long will the data last in this study?

The record acquired from the filming will be preserved indefinitely unless you request otherwise. Even after the interaction has been videotaped, you have the right to revoke your agreement to participate in this study and to remove your data from inclusion.

What are the risks of this study?

Because of the nature of the data being gathered (i.e., video-recordings), it may not be possible to conceal your identity as a participant. There is a potential risk that people known or unknown to you will formulate negative opinions of you or your behavior on the basis of their viewing of these data. We have three safeguards against this:

1. You can request that taping be stopped at any point during the activity, thereby preventing a record from ever being produced. In Week 2, you will be controlling the recording process.

2. You will be able to review the recordings we wish to use. If you wish the Principal Investigator will have particular segments containing your likeness deleted or erased.

3. As discussed in the section on confidentiality below, all additional data (i.e. names, addresses and any other revealing information) will be confidential.

Are there benefits to taking part in this study?

There are no direct benefits to you personally for participating in this study. The primary benefits from this work are for the advancement of scientific understanding of social processes and transportation problems. The availability of these data may lead to improvements in the social sciences and others areas of transport research and development of which you or others may be a direct or indirect beneficiary. If you are interested in these scientific outcomes, and we hope you will be, arrangements can be made with the Eric Laurier (the Principal Investigator) to be sent copies of your digitized files of the interaction, along with transcripts, analyses, and articles as they become available.

Will I receive any payment or other monetary benefits?

---

60 The Economic and Social Research Council, Polaris Avenue, Swindon, UK.

You will receive no payment for being recorded. You will however receive expenses for travel and food costs associated with your participation in the project on completion of your participation (we do not anticipate this exceeding £150 per vehicle). The ethnographic and video data will not be used by any member of the project team for commercial purposes. Therefore, you should not expect any royalties or payments from the research project in the future. We rely on and appreciate the goodwill of our participants in this project. As noted above we will offer you copies of your edited recording on DVD, CD or VHS.

What about confidentiality?

The nature of the data being gathered precludes completely concealing your identity as a participant. Any researcher who happens to know you personally may be able to identify you from the videotape. If this occurs, and you are unhappy about it, we will refrain from making further use of these recordings. In addition, we will keep any additional information that we have about you confidential through the use of pseudonyms. The original recordings will not be made available for any purposes outside of research activities.

What are my rights as a participant?

Taking part in this study is voluntary. You may choose not to take part or, subsequently, cease participation at any time.

Can I learn more about the project?

If you wish to know more you are welcome to contact Eric Laurier at the above address. There are also some websites that may be of interest to you:

The Habitable Cars project website

http://web.ges.gla.ac.uk/~elaurier/habitable_cars/

The Economic and Social Research Council

http://www.esrc.ac.uk/

If you have any other ethical concerns regarding the project then please contact: The Chair, Ethics Committee, Institute of Geography, University of Edinburgh, EH8 9XP.

Project Title : **Habitable Cars: the organisation of collective private transport**

Institute of Geography
School of Geosciences
University of Edinburgh

Research participant name : _____

As part of this project a video recording has been made of you while you participated in the research. We would like you to indicate below what uses of these records you are willing to consent to. This is completely up to you and you may withdraw from the research at this point if you wish. We will only use the records in ways that you agree to. In any use of these records, names will not be identified. *Please read the accompanying FAQ sheet before filling out this form.*

1. The records can be studied by the research team for use in the research project.

_____

[Please tick to indicate those permissions you wish to give]

2. The records can be shown to other participants in the research project.

_____

3. The records can be used for scientific publications.

_____

4. The records can be kept in an archive for other researchers.

_____

5. The records can be used by other researchers.

_____

6. The records can be shown at meetings of scientists interested in the study of community and social interaction

_____

7. The records can be shown in classrooms to students.

_____

8. The records can be shown in public presentations to non-scientific groups.

_____

9. The records can be used on television and radio.

_____

10. The records can be used on the research project website.

_____

I have read the above description and give my consent for the use of the records as indicated above.

Date _____

Signature _____

Signature of Guardian, if Applicable _____

Appendix 3: Transcription Notation

Symbol	Use
[text]	Start and end points of overlapping speech.
=	Break and subsequent continuation of a single interrupted utterance.
(seconds)	Time, in seconds, of a pause in speech
(.)	Brief pause, usually less than 0.2 seconds.
↓	Falling pitch.
↑	Rising pitch.
,	Temporary rise or fall in intonation.
-	Abrupt halt or interruption in utterance.
>text<	Enclosed speech was delivered more rapidly than usual for the speaker.
<text>	Enclosed speech was delivered more slowly than usual for the speaker.
°	Reduced volume speech.
ALL CAPS	Increased volume speech.
underline	Emphasis or stress in speech.
:::	Prolongation of an utterance.
(hhh)	Audible exhalation
(.hhh)	Audible inhalation
(text)	Unclear transcription.
((text))	Non-verbal activity.

Based on Jefferson, G. (1984) 'Transcription Notation' in: Atkinson, J. and Heritage, J. (eds), *Structures of Social Interaction*, Cambridge: Cambridge University Press.