**Supporting qualitative research in the humanities and social sciences: using the Mass Observation Archive**

**Fiona Courage and Jane Harvell**

**Introduction**

The availability of new technologies and tools for research in humanities and social sciences has changed the ways information professionals can support the use of qualitative data in humanities and social sciences. Recent years have seen a large increase in the number of digitization and metadata creation projects undertaken by libraries and archives across the world, underpinned by a firm acknowledgement from the research community that these resources are required to enhance and support work in various subject areas. Researchers are encouraged to use technologies to create cross-disciplinary and cross-institutional collaborations in their work; and by easing accessibility to qualitative data resources, we can support these initiatives as well as encourage the use of our unique and valuable resources.

For many years the Mass Observation Archive (www.massobs.org.uk/index.htm ), based at the University of Sussex, has provided researchers with a vast collection of qualitative data on many subject themes. Over the last ten years, the Archive has been involved in various digitization and metadata projects that use technology to increase accessibility to the collection. We shall look at the use of the Mass Observation Archive, offering a case study that examines both the common and the different elements of user support required by social sciences and humanities disciplines and how this might also be used to initiate and support collaborative research.

**Mass Observation**

Mass Observation was established in 1937 as a social observation project in which people around the country were recruited to become what Mass Observation’s founders described as ‘the cameras with which we are trying to photograph contemporary life’ (1)), constructing an organization that used a combination of ethnographic survey and reflective personal writing project to record everyday life in Britain. Between 1937 and the mid 1950s, over 2000 members of the public contributed to the National Panel of volunteer writers, sending in diary accounts of their daily lives and responding to monthly open ended questionnaires designed to elicit personal and subjective accounts of opinion and experience in contemporary Britain.

The purpose of all this activity was to give the opportunity for ordinary people to make their views heard, and provide access to these views to researchers from all fields. Initially various publications, journal articles, newspaper reports and broadcasts resulted from some of the findings, but ultimately much of the collected data was never used. The archive of material amassed in this initial project was eventually brought to the University of Sussex to be opened as a public access archive thereby fulfilling the original intention of making the information available to all who could use it.

**The Mass Observation Project**

The availability of the original project material inspired the establishment of a second phase of data collection beginning in 1981 under the direction of Professor David Pocock and Dorothy Sheridan. The concept of a National Panel of Volunteer writers was reawakened and a new batch of volunteers were recruited to respond to questionnaires that reflect on and record late 20th century life. The Project (www.massobs.org.uk/mass\_observation\_project.html) has continued unbroken since 1981, issuing three questionnaires or *Directives’* a year which normally deal with 3 different themes. Currently over 300 themes have been covered spanning themes as diverse as General Elections to Gardening, reactions to 9/11 to hair and hairdressing. The themes are often prompted by world events and current affairs, but over 30% have been commissioned by researchers who have opted to use the Mass Observation as one of their data sources.

The panel size averages out at around 500 members, some responding to only one or two Directives, whilst others have contributed for over 30 years. This provides a huge potential for longitudinal qualitative research; a data set of case studies reaching back years if not decades for researchers to access. Responses vary from one or two pages of writing to many pages of narrative, sometimes including photographs or ephemera to illustrate the responses. The unifying factor for these diverse responses is the qualitative nature of the material.

The Project often attracts comment that the panel is not representative of modern day Britain and indeed over the years there has been a slide towards the natural attraction of older women to contribute. Various attempts have been made to redress this since 1981 including the introduction of acceptance criteria in 2004, in the Project’s first significant attempt to recruit writers from areas of the population who were previously under represented. Mass Observation does not seek to establish social classification or ethnic background for its contributors, but it is likely that ethnic minorities are under-represented, and there seems to be a higher representation of urban dwellers.

The former Director of the Mass Observation Archive, Professor Dorothy Sheridan, believes that much of the discomfort experienced about using self- selected groups such as the Mass Observation panel stems from ‘a common belief about what constitutes *prop*er or scientific social research’ (2) Professor Sheridan illustrates how different disciplines encounter this type of data, in particular historians for whom:

‘…such material is a delight and a challenge because it may be all we have left of a particular life and time: a crucial part of the scholarly task is to establish the relationship between what has survived and its historical moment, that is, how “representative” can we take it to be and of what.’ (2)

Unlike many other data sets designed and produced for specific projects, the data collected by the Mass Observation Project is available for any researcher to use as soon as it becomes available. Even in instances where a researcher has commissioned a theme, they do not have exclusive access to the material collected. As such there are opportunities for different disciplinary interpretations for a single set of data. Whilst this certainly offers the possibility of exciting synergies and new opportunities, at the same time it places more importance on our role as data collectors, data controllers, interpreters and ethical advisors in our support for researchers.

**Researcher use of Mass Observation**

The Mass Observation Archive receives around 800 visits per year comprising mainly students and academics, the majority of which come from humanities backgrounds. The majority of these visits are made to look at the material collected largely between 1937 and 1950s, and of these researchers the main use is by humanities scholars, in particular historians. Similarly those looking at the post-1981 Mass Observation Project are largely from humanities-based disciplines with history researchers again predominating. This is in direct contrast to the profile of the researchers who commission new Directives from Mass Observation and are therefore seeking ‘contemporary’ data. Of those Directives that have been commissioned since 1991, eleven disciplines are represented: 78% of the Directives commissioned have been done so by social science researchers, 16% humanities and 6% sciences. These two distinct types of usage illustrate how a single resource can be used in two different ways by different disciplines: humanities scholars tend to use the collection as a historical archive or primary source evidence whilst social scientists use it as a way of assembling a contemporary data set.

**Supporting research use through digitization**

Perhaps the greatest change in the way user support has been delivered to researchers in recent years is the rise in the availability of online resources including digital facsimiles of archival materials. Over a decade ago the benefits of using digitized materials were becoming more apparent, providing easy availability at a location and time convenient to the researcher, and the ability to perform searches and make copies of material that could be difficult when using microform or originals (3). Advances in technology have served to make these factors even more beneficial to the researcher with improvements in tools such as Optical Character Recognition providing enhanced searchability and legibility for many resources. Despite these advances, it has been acknowledged that many humanities scholars, and to a lesser extent social scientists are less advanced than scientists in their adoption of digital resources as a primary tool for working with archival collections(4) identifies the need for analytical tools and services to become more sophisticated and transparent for the humanities community to use; these scholars in particular are often in search of nice physical distinctions sometimes lost in the transition to digital format, requiring a high dependence on metadata to identify and compare variants. Creation of this type of metadata is naturally linked to the role of librarian and archivist whose cataloguing and curatorial skills make enable them to develop transparent and easy to use digital resources that will act as a strong surrogate for original hard copy primary sources.

The Mass Observation Archive illustrates how a primary source resource primarily used by humanities researchers has been digitized and provided with tools and data designed in conjunction with archive staff to support researchers. In 2006 the Mass Observation Archive embarked on a partnership with the commercial publishing company, Adam Matthew Digital (www.amdigital.co.uk), to create a digital resource, Mass Observation Online (www.amdigital.co.uk/Collections/Mass-Observation-Online.aspx). Throughout this project sequential tranches of the Archive from 1937 -1950s have been digitized and released for sale around the world, with over 80% of purchasers being outside the UK. By 2015, the entire holdings of this first phase Mass Observation (1937-1950s) will have been published in this way.

The publisher’s statistics indicate that *Mass Observation* *Online* has received increasingly heavy use since publication, with over 19,500 unique visits being made by users from institutions that have purchased the resource worldwide between 2008 and 2011. Despite the extent of material already available electronically, the statistics for visits by UK and international visitors to access the Archive at the University of Sussex have remained stable. This would seem to indicate that far from undermining the use of the physical Archive, digitization has merely widened accessibility and thus increased usage.

Digitizing the collection has also brought a level of flexibility in the way researchers might use the Archive not offered by the physical collection. As noted in the Research Information Network (RIN) Report (2007), scholars were found to engage with ranges of resources and technologies, mixing digital with hard copy and being prepared to adopt technologies to improve their current practices:

‘They have become used to managing digital resources and this has freed them to access and use information which is in locations far from its source. Moreover, the very nature of digital technologies has enabled researchers to create and assemble information in new ways in the course of their research, presenting new issues to them…’ (5)

Publishers and creators of these digital resources must understand these needs feed the potential and value of digitizing primary source material. As part of the creation of *MO Online* specific tools were developed in response to researchers’ expressed demands. One example is a mapping tool that allows researchers to locate quickly and easily diarists in the same region of the UK without needing to interrogate a database.

For humanities researchers the ability to follow themes and individuals through the collection was important. The large, hard copy collection of diaries have no subject index and are held in chronological order so that a researcher looking for a specific subject needs to read through them in search of relevant material. Equally, if a researcher wants to track an individual diarist, they need to order a box of material for each month the diarist wrote – potentially tens of boxes need to be gone through. Digitization provided the opportunity for researchers to create their own indexes, whilst a with a few clicks of a mouse a researchers are now able to pull up a listing of individual diarists, and then trawl through all they have written rather than having to plough through diary entries of hundreds of other diarists. In some cases optically character read (OCR) technology has been used to create searchable rich text.

However it is worth noting that one of the key findings of *Information Practices in the Humanities* (6*)* was that although

‘…they are reluctant on occasion to consult texts that require a trip to a distant library or archive. Nevertheless, none of the participants is yet ready to abandon print and manuscript resources in favour of digital ones. Rather, they engage with a range of resources and technologies, moving seamlessly between them.’

This report concluded that ‘Such behaviours are likely to persist for some time’ (6)

**The curatorial voice in a digital world**

The use of digital proxies has strong implications for user support. Unlike personal visitors to the Archive, there is no member of staff on hand to explain the intricacies of catalogues or the context of collection holdings. So it is important to ensure that the digital resource replicates the ‘curatorial voice’ for the researcher. The need for an effective ‘curatorial voice’ highlights the importance of collaboration between curator and publisher to develop a coherent set of guidelines not only to help the researcher to understand the resource, but that also to offer them as profound a user experience as they might experience in person. Working closely with Mass Observation, Adam Matthew Digital was very careful to offer this voice and context, adding value for the researchers in their use of *MO Online*:

‘When creating our digital resources, we always strive to create a sense of context, which is absolutely essential if users are to research effectively online. Without this, users may find themselves unable to understand or navigate the archival content. This sense of context can be created by careful consideration of how to organize the digital material (in MO Online's case, we tried to be as consistent as possible with the physical arrangement of the archives); but also through use of secondary resources aimed at different user levels, such as essays designed for undergraduates or for researchers, which can recreate the invaluable experience of being shown round an archive by a 'real' curator. Even the front end design of the website can help build an 'atmosphere' that suggests ways to approach the source material. By these means we hope that users can be guided to understanding what the digital archives offers them’

(Martha Fogg, Senior Development Director, Adam Matthew Digital)

The sympathetically re-created ‘curatorial voice’ available on *MO Online* means that the electronic resource goes beyond the mere provision of a corpus of digitized material, but also provides context and understanding for users of this large and unusual collection.

All these features have ensured that the digitized resource can become an invaluable part of the way the collection is used, both remotely and *in situ*, thereby creating added value to the material object itself.

**Mass Observation as data set collection/collector**

As stated earlier, humanities scholars tend to use the collection as a historical archive or primary source evidence, whilst social scientists use it as a way of collecting a contemporary data set. Supporting this notion is the fact that social scientists, in particular sociologists, dominate the numbers of researchers using Directives as a data set for their research commissioning 78% of Directives issued since 1991. The questionnaire is currently drafted by Mass Observation staff in collaboration with the commissioning academic, and sent to the Panel as part of a regular mail out. Responses are collated by Archive staff and made available in the reading rooms to any researcher within 3 months of the mail out date. This material has not yet been digitized and made publicly available.

An interesting usage pattern of has evolved among the social scientists who commission these Directives. Unlike many of the humanities scholars who visit the reading rooms in person and work through boxes of hard copy material from various themes, social scientists often request copies of the entire set of responses to a single Directive theme. These responses are sometimes used as case studies for qualitative analysis, occasionally using the potential for longitudinal study of individuals to track changes in habit and opinion by looking at their responses to similar themes over a period of time (7) Coding the responses is also a common practice, sometimes using specialized software such Atlas (www.atlasti.com) (qualitative data analysis and research software), sometimes simply by cutting and pasting into tables and applying specific word searches to spot and analyse patterns.

Whilst new Directives are often used as data sets, within a short space of time they can also be used as primary source material for other disciplines. As information professionals, it is our duty to make the information available in a way that will best serve the different needs of social science and humanities researchers alongside other disciplines that may use the material. So it is important for this material to be catalogued and described in ways that can cross these boundaries and enhance the potential for future discovery and collaboration.

**Tools for support**

‘Creating high quality metadata and data documentation can help the user interpret raw data sources*.’* (8)

The nature of the information professional’s role is to provide support for using resources, in its most basic form – letting the user know what is available. Cataloguing underpins the accessibility of a collection with metadata enhancing its usability by allowing the user to interrogate and manipulate data as they need. Whilst an international standard for archival description (ISAD[G]) (www.icacds.org.uk/ eng/standards.htm) has been used to catalogue the Mass Observation Archive and make it available through the *CALM* (www.axiell.co.uk/calm)archival management system accessible by users, we have consistently sought ways to augment this information with metadata to make it more searchable. As Louise Corti, Associate Director at the UK Data Archive ([www.data-archive.ac.uk](http://www.data-archive.ac.uk)) notes:

‘Users of qualitative data want easy access to data and they want more than just raw data. *Enhancing* collections so they can be used more easily and effectively should be central to an archives mission. It is never just about preserving original research documents.’ (9)

In 2009-2010 a retrospective cataloguing project listed each individual response to Mass Observation Directives since 1981 onto the CALM database. Before this was done, the number of responses to a Directive would have been unknown to a researcher. The project has enabled the provision of an entry for each response to every Directive containing metadata and thus offers the researcher details on the item and the writer. Using this metadata, each individual writer can now also be linked to their list of responses, a process that previously could be undertaken only by archive staff using card catalogues and so would be lengthy and costly. Researchers are now able to identify resources they would like to access, follow up lines of enquiry and pin point potential case studies for their work.

In 2010 another project designed to enhance the accessibility of information on the Mass Observation Archive was undertaken. Funded under the *JISC Infrastructure for Resources Discovery Programme*, (www.jisc.ac.uk/whatwedo/programmes/inf11/ infrastructureforresourcediscovery.aspx) the *Sussex Archive Linked Data Application* (SALDA*)* (<http://blogs.sussex.ac.uk/salda/about>) project extracted metadata from 23000 records from the Mass Observation Archive catalogues on the *CALM* system. This data was converted into Linked Data (<http://linkeddata.org>) and made publicly available using an Open Licence via *Talis* Platform (http://www.talis.com/platform). Alongside creating a set of Linked Data from a single collection, the project provided the Archive staff with valuable skills in understanding the methodologies involved in transforming data in this way, as well as opening up possibilities for wider joining wider resource discovery networks in the future.

The concept of sharing data and making it available through other portals and institutions is valuable in terms of heightening the visibility of collections and making them useful to researchers. A current JISC funded project, Observing the 1980s(http://blogs.sussex.ac.uk/observingthe80s/about), is being undertaken to create raw and cooked Open Educational Resources (OERs) from material collected as part of the Mass Observation Project in the 1980s and sound recordings held in the British Library. These OERs are primarily aimed at learning and teaching, but an off-shoot of this project is the creation of standardized summaries of items with which they can be identified and classified. This set of summaries, or enhanced metadata will be deposited at Qualidata ([www.esds.ac.uk](http://www.esds.ac.uk)) and thus become openly available to an even wider set of researchers.

**Serving changing research needs - collaboration**

‘A key change in humanities research over the past 10-15 years has been the growth of formal and systematic collaboration between researchers. This is a response in part to new funding opportunities, but also to the possibilities opened up by new technology’ (6)

Researchers are working with new technologies in collaborative ways to use resources in increasingly diverse ways. Information professionals can support research by providing environments that are conducive for interaction between scholars, encouraging them to develop synergies over specific resources.

In 2006 Mass Observation set up a JISCMail ([www.jiscmail.ac.uk/](http://www.jiscmail.ac.uk/)) list to encourage researchers and those interested in Mass Observation to enter into discussions regarding the collection, and to share experiences of using the material in their research. The establishment of this list reflected Mass Observation’s original objective of mass collaboration with the aim of gathering information and making it available to all who might usefully use it.

Other collaborative research projects include the University of Brighton- funded collaborative research network, *Methodological Innovations: Using Mass Observation*, ([www.brighton.ac.uk/sass/research/massobservation](http://www.brighton.ac.uk/sass/research/massobservation)) a vibrant collective of researchers from arts, humanities and social science backgrounds within the University working together to create opportunities to share research works and to establish new areas of collaborative research.

**It’s about people as well as resources and technology!**

Being open to organizing and participating in conferences, seminar series and events means that we can highlight the potential for using collections as a research resource as well as illustrating the range of support that we can give. Taking advantage of its 75th anniversary year, Mass Observation has organized a public lecture series and an academic conference ([www.massobs.org.uk/conference.htm](http://www.massobs.org.uk/conference.htm)) drawing on speakers who have used the collection in a variety of ways. In each event, there is the opportunity for researchers to network and find out how others have used the Archive. These occasions which bring scholars, the general public and Mass Observers themselves together are building on experience gained from previous successful events such as the Conference organized by Mass Observation in 2007. They provide the opportunity for researchers to publicize their research work and interests within a ‘market place’ of posters and displays. Events such as these create tremendous possibilities for networking and identifying potential partners for future collaboration.

**Curating Ethics**

Finally, Mass Observation acts as collector, guardian and gatekeeper of data collected under its aegis. Responses are anonymised before being given to researchers and any interaction with the panel is handled by Mass Observation staff. As curators, we are bound to protect the subjects of the material, as well as we protect the material itself.

In a paper on the challenges of web access to archival oral history in Britain, the Oral History Curator at the British Library, Dr Rob Perks, discusses how keepers of materials exercise control over access. He argues that changing the way researchers access data (in this case via the internet) transforms the guardianship relationship which we as curators have with the material. He claims that this can make us feel uncomfortable as ‘we fear a marginalization of our own role as conduit, explainer and interpreter.’ (10)

Throughout this process it is vital that as curators we communicate the reasons for our actions to researchers to develop the process of trust for both the organization and the data collected. We are able to support the researchers in developing and research ethics suited to their particular discipline’s use of qualitative data, both generally, and more specifically, Mass Observation data , in their work.

**Conclusion**

Success for the Archive comes with strong academic involvement both in terms of creating new data sets and providing access to it as the historical archive as a whole. Responsible storage and accessibility of good quality metadata is vital to ensuring the Mass Observation materials are easily discoverable and open. By working closely with academics and information profession colleagues we have been able to create projects and tools that enable researchers from the social sciences, humanities and other disciplines to use and manipulate this vast collection of unique qualitative data. It is important that we present an Archive that researchers can trust as an established organization with a longstanding academic reputation. This chimes nicely with the assertion in the RIN report:

‘..it seems more likely that concerns over the vulnerability of socially-created valuable information will send researchers into the arms of those they trust – on curation and preservation issues at the very least’ (5)

**References**

1. Harrisson, T.H. and Madge, C. (1938) *First Year’s Work 1937-1938*. Lindsay Drummond.
2. Sheridan, D. (1996) *‘Damned anecdotes and dangerous confabulations’: Mass Observation as Life History* Mass Observation Occasional Paper Series no. 7
3. Duff, W.M. and Cherry, J.M. (2000) Use of historical documents in a digital world: comparisons with original materials and microfiche. *Information Research,* **6**(1).
4. Borgman, C. (2009) The digital future is now: a call to action for the Humanities, *DHQ: Digital Humanities Quarterly* **3**(4)
5. Research Information Network (2007) *Researchers’ use of academic libraries and their services.*
6. Research Information Network (2011) *Reinventing research? Information practices in the humanities.*
7. Busby, H. (2000). Health, sickness and the work ethic. *Mass Observation Occasional Paper No. 11*. 7.
8. Corti, L (2005) User support, *Forum: Qualitative Social Research* **6**(2), Art 41.
9. Corti, L. (2011) The European Landscape of Qualitative Social Research Archives: Methodological and Practical Issues, *Forum: Qualitative Social Research* **12**(3), Art 11.
10. Perks, R, (2009). The challenges of web access to archival oral history in Britain