



Learning to manage and share data: jump-starting the research methods curriculum

Louise Corti & Veerle Van den Eynden

To cite this article: Louise Corti & Veerle Van den Eynden (2015) Learning to manage and share data: jump-starting the research methods curriculum, *International Journal of Social Research Methodology*, 18:5, 545-559, DOI: [10.1080/13645579.2015.1062627](https://doi.org/10.1080/13645579.2015.1062627)

To link to this article: <http://dx.doi.org/10.1080/13645579.2015.1062627>



© 2015 The Author(s). Published by Taylor & Francis



Published online: 25 Jul 2015.



Submit your article to this journal [↗](#)



Article views: 292



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

Learning to manage and share data: jump-starting the research methods curriculum

Louise Corti* and Veerle Van den Eynden

UK Data Archive, University of Essex, Colchester, Essex CO43SQ, UK

(Received 31 October 2014; accepted 29 March 2015)

Researchers' responsibilities towards their research data are changing across all domains of social scientific endeavour. Government, funders and publishers expect greater transparency of, open access to, and re-use of research data, and fears over data loss call for more robust information security practices. Researchers must develop, enhance and professionalise their research data management skills to meet these challenges and to deal with a rapidly changing data sharing environment. This paper sets out how we have contributed to jump-starting the research methods training curriculum in this field by translating high-level needs into practical guidance and training activities. Our pedagogical approach involves applicable, easy-to-digest, modules based on best practice guidance for managing and sharing research data. In line with recent findings on successful practices in methods teaching, we work on the principle of embedding grounded learning activities within existing narratives of research design and implementation.

Keywords: data management; data sharing; teaching approaches; research methods curriculum; transferrable skills

Introduction

Across the disciplinary spectrum, researchers' responsibilities towards their own research data and their use of others' data are changing. Open access is increasingly mandated for publicly-funded research data, governments demand transparency in research they support, and publishers ask for evidence of data that underpin findings. The economic climate also requires more re-use of data. In turn, concerns about data loss call for more robust information security practices. Much of the responsibility for the management of research data is placed upon researchers who need to improve, enhance and professionalise their skills to meet the challenge of producing the highest quality research data for publication, sharing and reuse in a responsible way. Since data form the cornerstone of empirical research, such data skills must be integrated with research methods teaching and learning. Novice researchers at undergraduate or postgraduate level may benefit from teaching that is integrated into degree curricula and sits naturally alongside narratives of standard methodological approaches. Experienced researchers are likely to need a greater degree of upskilling, to plug gaps in current knowledge and refresh or update their knowledge in response to rapid changes in technology or legislation relating to the governance of

*Corresponding author. Email: corti@essex.ac.uk

research data. Irrespective of the level of a researcher's expertise, these skills can help maximise the impact of research and foster greater appreciation of the use of secondary sources. However, guidance and training for gaining data management and sharing skills are noticeably absent from many social science research methods curricula, with the exception of data handling covered within statistical analytical techniques. Data management skills are equally important for qualitative researchers. A culture change is therefore needed and pedagogy forms part of that change.

In this paper we show how research methods training can be enhanced by integrating practical guidance and training activities into research methods instructional literature and teaching. In this approach, upskilling can be achieved by interweaving modular practice-based instruction in data management with traditional methods learning that focuses on research design, fieldwork and analysis. As we move into a growing data landscape with novel forms of data coming on stream, these skills represent a step change towards the next level of competence required for data mining and data science.

Our pedagogical approach involves up-to-date, easy-to-digest information about practices for managing and sharing a variety of types of research data, presented in a modular structure. In line with recent findings on successful practices in methods teaching, we work on the principle of embedding grounded learning activities within the narrative of research design and implementation (Buckley, Brown, Thomson, Olsen, & Carter, 2015). Effective learning of data skills within social science research methods teaching was found by Kilburn, Nind, and Wiles (2014a) to be achieved through three means: active learning by making processes visible; directly experiencing methods; and critical reflection on practice. Providing building blocks that create a solid empirically-based research knowledge base to underpin the creation and use of digital data in the social sciences is at the heart of our approach. A modular structure is also beneficial because it enables the coverage of topics to be tailored towards multi- or non-disciplinary learners. Case studies, practical activities, exercises and group discussions help support effective learning and application of the core concepts. Success of the approach has been demonstrated through a grant from the UK Economic and Social Research Council's (ESRC) Research Development Initiative to refine and test the teaching materials and by integration of content and pedagogy into more recent publication of an internationally-appealing handbook on the topic (Corti, Van den Eynden, Bishop, & Woollard, 2014).

Challenges for data management pedagogy

Emerging agendas for data sharing

In the past 5 years, significant funds have been made available to open up a variety of data for research purposes, such as government administrative data, data underlying published research outputs and data from commercial sources. The rise of the 'big data' agenda has also pushed to the fore the issue of how to look after and manage data to benefit interdisciplinary investigation of large and complex data. In the UK, funding is being dedicated to secondary analysis research, with the intention of maximising opportunities for the analysis of data from large-scale government investments such as national surveys and cohort studies (ESRC, 2012). With a decrease in funding for research generating new data, it is important for researchers to explore the optimal use of existing data. Investments that seek to advance data-related skills are also being driven by research councils' policies on open access

(Research Councils UK [RCUK], 2012) and by the desire to enhance quantitative skills in social science education (British Academy, 2014). These sets of activities would appear to benefit greatly from alignment around shareable data as a focus for training in the collection, management, and analysis of long-term data resources.

Skills involved in data management

While data management skills come under the broad banner of data handling, they fit in at all stages of the research process. Figure 1 shows likely intervention points in the data lifecycle of data management type activities that should be considered in any research design and throughout the research process.

Based on the UK Data Archive's decennia of work supporting and advising ESRC grant holders about data archiving, sharing, and reuse, within the framework of the ESRC research data policy, this framework was translated into guidance that followed the logic of the data lifecycle. Guidance was iterated, enhanced and expanded into detailed best practice guidance for researchers and accompanying teaching materials and further development resulted in an extensive handbook on best practices for managing and sharing research data (Corti et al., 2014). The content includes:

- the importance of sharing data;
- consent, confidentiality and ethics of data sharing;
- rights associated with the use of existing data;
- describing, contextualising, and documenting research data;
- data formats and software;
- storage, back-up and security in response to the fragility of digital data;
- publishing data in a sustainable way and enabling future citation.

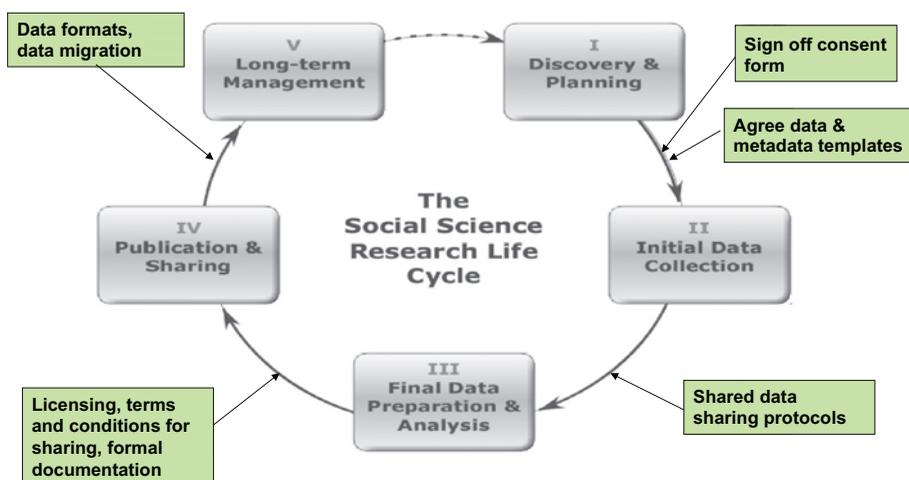


Figure 1. Intervention points in the research data lifecycle. Based on Green and Gutman (2006).

Proficiency in these topics is expected for those creating new data during the course of research and those undertaking secondary analysis. Enhanced data skills become essential as we move towards a new data landscape where more digital data are increasingly available for use, including new forms of data such as social media discussions, administrative data, commercial customer data and epidemiological data. Such novel data sources are already entering the social sciences via new data centres dedicated to the challenges of solving research challenges using ‘big data’ (ESRC, 2013). In the future we are likely to see increased analysis of large and varied data sources in the pursuit of impactful findings.

Data handling in methods teaching: assessing gaps

Data handling methods have tended to be more widely acknowledged in disciplines where large-scale secondary data sources are routinely used. Our own experience tells us that, for example, geographers and historians tend to have good data skills and are familiar with reusing and sharing research data. Secondary analysis of survey data is routinely taught, and here data management skills such as quality assurance measures, metadata and annotation of data, formatting and modes of organisation are very much part of research and learning activities. Assessment of data created by someone else is a useful way to gain insight into what it means to prepare data for sharing and future use. In turn, economists may have good technical data skills, but sometimes lack appreciation of ethical issues surrounding data obtained from human participants (Blomfield, 2012). Other research domains may suffer from a tendency to exclude technical components from methods courses. For instance, students of sociology or linguistics may only encounter data handling in the context of collecting data for their own research projects.

Despite reviews of postgraduate training guidelines and surveys of senior academics and researchers over the past 10 years, the UK still fails to produce enough quantitatively proficient or data-savvy graduates in some areas of social sciences (MacInnes, 2012). From an international perspective, the quantitative skills deficit was also recognised as a major concern in the 2010 *International Benchmarking Review of Sociology* (ESRC, BSA, & HAPS, 2010). This recognition by research funders and professional societies propagated a welcomed flurry of activity (British Academy, 2012; Nuffield Foundation, 2012; Royal Statistical Society, 2014; Williams, Payne, Hodgkinson, & Poole, 2008). Concrete investments have followed, with some £20million invested in the UK’s recent Q-Step initiative that aims to achieve a life-course approach to building quantitative methods skills (Nuffield Foundation, 2013). The UK’s ESRC National Centre for Research Methods (NCRM) is also supported to initiate and promote targeted and advanced methods training across the methods spectrum (NCRM, 2014). This drive to embed methods in real social science questions and engage in practice-based learning rings well with our own approach. However, it would be beneficial for these programmes to explicitly consider the longer-term value and re-use of data.

Of most interest to the thrust of this paper were the findings emerging from the 2012 Higher Education Academy Social Sciences learning and teaching summit, focusing on teaching research methods in the social sciences. These included innovative ways of improving applied research methods, for example by gaining experience outside of the academic environment and undertaking projects in the work place (Hamilton, Gossman, & Southern, 2014; Kirton, Campbell, & Hardwick,

2014). These offer working models of applied methods in practice. Others have identified real life problems and task-oriented group work as a means of engaging reluctant students with research methods teaching (Leston-Bandera, 2013).

We have seen some progress in the uptake of secondary analysis, particularly for qualitative researchers. The outreach work done by the UK Data Service and its predecessors in providing teaching resources and running training courses encourage uptake of secondary analyses for qualitative data (Bishop, 2012). This included collating case studies from academics who use secondary data in their own teaching (UK Data Service, 2014c). A recent review of research methods learning and teaching also recognised the opportunities of using existing data as a learning tool in methods teaching – either working with own data that learners bring to a session, or using teachers’ data or teaching datasets, the latter being easier to manage (Kilburn, Nind, & Wiles, 2014b). Across the curricula, though, there remains work to be done on streamlining the portfolio of data management skills within research methods teaching.

The supporting research methods literature

The recent research methods literature continues to comprise publications that predominantly focus on research design, data collection in the field and data analysis, although the availability of instructional books on secondary analysis techniques is growing (Smith, 2008; Trzesniewski, Donnellamn, & Lucas, 2011; Vartanian, 2010). However, the social science literature still lacks texts that cover the core skills required to support good data management and sharing. Some of these skills are inherently taken for granted as good research practice. Yet issues such as file format choices, storage, backup and security of research data, and documentary descriptions of data, are crucial in creating high quality data. So too are ethical and legal considerations in data use and stewardship – both within and beyond the life of project. In particular, qualitative researchers may find few written sources on managing digital data for sharing. Authors of new instructional methods books should be encouraged to include such topics.

Upskilling for novel forms of data

In the ‘big data’ era, researchers see the benefits of working with novel and more complex data sources, often across disciplinary boundaries. With this, there is potential for deficits in confidence or skills for researchers to become even greater. While some disciplines may feel comfortable running algorithms on big data, many social scientists will want to take smaller samples and analyse these in more familiar ways. Examples are seen in the use of social media data such as Twitter feeds, where a discourse analyst may wish to select tweets from a particular social group, geographical context, or pertaining to a specific issue.

Good metadata and quality assurance skills are needed for retrieving, assessing, manipulating, and analysing these big data extracts; skills that fall under the rubric of data management. Courses aimed at non-technical audiences are emerging on the use of social media data, such as network analysis and data mining. These courses will also introduce skills on data extraction using Application Programming Interfaces and assessing the quality of new forms of data by means of statistical techniques (University of Michigan, 2014; University of Southampton, 2014).

For instance, research on smart cities demands much greater appreciation of data modelling, which is best achieved through cross-disciplinary methodological collaboration (Urban Big Data Centre, 2014).

For the qualitative researcher wishing to explore the potential of online digital textual sources – be those social media data or other source – the upskilling needed to handle, extract and analyse such data may be significant. Basic core data management skills can bring researchers some steps closer to technical proficiency. For instance, a successful outcome might be for a biographical researcher to run algorithms on large volumes of rich text relevant to their area of study.

Despite the trends towards exploiting ‘big data’, users of ‘smaller’ data should not be left by the wayside. While small scale research will remain critical, researchers who perhaps consider that big data pushes them out of their comfort zone can benefit from upskilling. Enabling them to feel comfortable with handling complex and new forms of research data, to be cognisant of longer-term sustainability of data, and to appreciate the art of accessing and re-use of secondary data sources are all significant steps towards this goal. This incremental knowledge sharing and capacity building will help the broader social science research community move towards appreciating big data.

Looking to the more data-intensive disciplines, the new data science courses aimed at computer or statistics students all take for granted some areas of data management skills with their modules mostly oriented towards data handling, quality assessment and analytics. However, it may be some time before undergraduate courses in traditional social science disciplines embrace such technical components. Even in our own institution, where data science courses have recently started up, the social sciences do not yet contribute to that teaching with students instead hailing from maths and computing programmes. The emergence of internet-based text as a data source is prompting qualitative social scientists to consider branching into new types of analysis, requiring more advanced computational skills which may benefit from structured training in the context of degree courses.

In summary, the UK’s apparent quantitative methods skills deficit, together with demands of the current research climate (openness, accountability, impact and value for money) and the need for computational skills in the era of big data, emphasise the importance of data management skills. To meet this need, methods curricula must evolve rapidly.

Applying solutions: data management skills, training, teaching and learning

Professional training in data management for researchers

Sharing data is becoming a vital part of research activity. This is especially true if we consider the training needs of professional researchers and their requirement for life-long learning to keep pace with the changing demands of social research and the UK research councils’ expectations for societal and economic impact (Research Councils UK, 2011).

Over the past 5 years we have seen a rise in the marketing of career development pathways and support for junior researchers. Vitae (2013), an international programme led by a UK charity and supported by UK research and higher educational funding, is dedicated to realising the potential of researchers through transforming their professional and career development. In their role of supporting the employability of doctoral researchers, Vitae concentrates on enhancing research skills.

However, a review of their 2008–2012 framework and concordat indicates that, while employability of postgraduates is high on their agenda, there is no mention of the need for data competence or management skills across the disciplines.

Social science libraries also play a role in offering support to students in gathering references for assignments and in some cases actively helping with statistical projects. In the US there has long existed the specific role of the data librarian, working to deliver access to data and even teaching appropriate techniques for statistical analysis. Evidence from Corti and Watkins (2004) suggested that in the US, where there is typically a heavier focus on quantitative methods training in social science courses, students needing guidance with statistical assignments turned to their local data librarians for assistance. Consequently, a movement of dedicated extra-curricular training by data librarians to support these skills has developed.

In the UK, the Jisc (the lead organisation that champions the use of digital technologies in UK education and research) invested in a 5 year Managing Research Data programme (2009–2013) to help develop and deliver data infrastructure, resources, tools and training materials to meet the increasing demands of research funders for good data management or sharing of research data in the higher education sector. This resulted in various UK institutions developing and investing in institutional research data policies, data repositories and data management training resources. Examples include the Universities of Edinburgh and Bristol, where materials were developed partially based on the UK Data Service training modules (UK Data Service, 2014b; University of Bristol, 2014; University of Edinburgh, 2014). The Digital Curation Centre provides an overview of data management training provided for researchers and data custodians, such as library and research support staff across UK institutions (DCC, 2013). The UK Data Service runs a national programme of training for researchers and research support staff (UK Data Service, 2014b, 2014c).

There are also an increasing number of training courses on the topic of digital preservation (now a discipline in its own right). Training in digital preservation is aimed at information and IT professionals, as well as archivists and curators of digital heritage and collections, rather than researchers exclusively. These seek to provide hands-on knowledge and skills involving the collection, selection, management, long-term preservation and accessibility of digital assets. Examples are the Digital Preservation Coalition training courses in the UK (Digital Preservation Coalition [DPC], 2014) and the Library of Congress's Digital Preservation Outreach and Education Programme in the US (Library of Congress, 2013). Some master programmes for archivists now cover useful aspects of data management, though are unlikely to be grounded in a narrative of social science research.

In summary, courses on data management for sharing are emerging in academic institutions in the UK and US, motivated by research funders' data policies. Europe is lagging behind, largely due to the absence of funder data policies and the drivers these bring.

A modular strategy for introducing data management skills

Doctoral training in methods has drastically changed in the last few years, with a move away from ad hoc supervisor–student training towards institutionally organised training, making use of short and professional development courses, hands-on sessions and small group work, plus case studies critique (Kilburn et al., 2014b; Kottmann, 2011; Sloan, 2013; Strayhorn, 2009).

Examples of teaching advanced quantitative methods using a modular approach can be found in the online courses for multi-level modelling provided by the Learning Environment for Multilevel Modelling (LEMMA) (Centre for Multilevel Modelling, 2014). LEMMA offers a set of graduated modules starting from an introduction to quantitative research progressing to multilevel modelling of continuous and binary data. Equally, In the US the Interuniversity Consortium for Political and Social Research's Data-Driven Learning Guides enhance teaching of core concepts making use of topics drawn from concepts that are included in introductory-level social science textbooks (Inter-University Consortium for Political and Social Research, 2009). Their guides are topic-based and pose research questions that could be explored using a particular data-set. An example is the Guide on *Age and Attitudes about the Rights of Homosexuals* which investigates trends in attitudes regarding the rights of homosexuals in the United States from the early 1990s to 2007, using the longitudinal Houston Area Survey. The module guides users to work through the questions, prompting them to recode the data, available via an online browser, to be able to answer substantive questions of interest. Commentary is provided taking the reader through analytical issues relevant to the data, such as inference and weighting, and provides a summary of key findings.

In developing research data management training materials on a range of modular topics, that can complement methods teaching in quantitative and qualitative research methods, we have formulated a flexible toolkit that can be incorporated into undergraduate, masters, and doctoral taught curricula. These modules can also be used as voluntary symposia, doctoral colloquia, or professional development sessions. While the skills require some degree of research methods knowledge, they remain accessible to second or third year undergraduate students.

Over the years, our training workshops on managing and sharing research data have developed to achieve effective learning of data skills, following the forms that are recognised in existing literature on research methods teaching: active learning by making processes visible; directly experiencing methods; and critical reflection on practice (Kilburn et al., 2014a). Active learning is realised through the use of numerous practical tasks, exercises and discussions during training sessions, such as hands-on lab exercises in encryption of sensitive data files, creation of metadata using existing software tools, anonymisation of interview transcripts and survey tables, and quizzes on data practices in collaborative research. Learning by doing or experiencing is applied where trainees develop their own data management plans throughout training sessions. Critical reflection on data practices is supported through group discussions of real-case data challenges, such as the ethical implications of sharing confidential data or copyright scenarios, or discussions around individual's own data requirements to explore shared challenges.

Our modules consist of theoretical presentations, hands-on exercises or scenarios and quizzes (with answers) to practice understanding of topics we consider essential in conducting research, efficient collaborative working, and undertaking informed secondary analysis. The modularised topics are: why and how to share research data; planning data management; documenting and contextualising data; formatting data; storing, backing-up and transferring data; ethics, consent and confidentiality; and copyright of data (Corti et al., 2014).

These modules were road-tested and fine-tuned into flexible training materials through use in over 75 hands-on workshops to audiences of doctoral students, researchers and research support staff over the period 2008–2013. Research methods

trainers can adapt and repurpose presentations and exercises to suit their own discipline, country, or relevant laws. The resources are downloadable via our website. Although the resources were primarily developed for social scientists, they are broad enough to be used by other disciplines including the natural sciences, arts and humanities.

Fitting modules to the audience

In running training sessions, we adapt programmes to suit our audience. For example, for a two-hour doctoral professionalisation workshop for social scientists, time would permit a presentation and an exercise on gaining consent for collecting, using and sharing qualitative data from structuring interviews in the form of well-documented and well-formatted transcripts. A half-day workshop might focus on several technical modules on formatting and organising, storing, and documenting data, as a continuous professional development course. The use of suggestive titles like *Are you looking after your research data properly?* help to attract audiences.

Some topics lend themselves better to facilitated group discussions, such as sharing experiences of managing data-rich projects. More technical areas, such as data storage or encryption, may benefit from relatable anecdotes of data disasters in presentations and the realisation of understanding through light-hearted quizzes, so as to keep learners engaged.

Here we provide four examples of how a specific audience can be taught the core skills within a set time frame. We discuss the most successful trajectory for delivery of these modules, when to condense or lengthen them, and what kind of activities work best in particular settings. The following workshop examples a two hour overview session to an intensive two-day workshop.

- (1) Two-hour taster session for students (second year undergraduates onwards) on *Looking after your research data*:
 - Introduction to data managing and sharing, planning data management: *presentation*.
 - Data storage, back-up, security, transmission, encryption, file sharing: *presentation, exercise, quiz and tools demonstration*.
 - Quality control, version control, formatting, describing and organising data: *presentation, Q&A and quiz*.

- (2) Half-day training for social science professionals/ethics committee members on: *Managing and sharing research data: legal and ethical issues*:
 - Data management and data sharing applied to research involving people as participants: legal and ethical aspects of data sharing (based on UK legislation and requirements for publically funded research) and data management planning: *presentation and Q&A*.
 - Developing consent agreements for obtaining informed consent from participants to share data: *presentation, exercise, group discussion*.
 - Dealing with confidential research information and anonymisation techniques to enable use and sharing of research data: *presentation, exercise, group discussion*.

- (3) One-day training for social science and humanities research students on *Going digital: Looking after and managing your digital research data*:
- Benefits of data management and data management planning: *presentation and discussion*.
 - Documenting your data: *presentation*.
 - Creating metadata for a data collection: *practical lab session with exercises*.
 - Storing data, data security, formatting, encrypting, organising data: *presentation, software demonstrations and Q&A*.
 - Encrypting files, backups, checksums: *practical lab session with exercises*.
 - Copyright and IPR: *presentation and scenarios for group discussion*.
 - Ethical and legal aspects – key messages for sharing data: *presentation and scenarios for group discussion*.
 - Data management planning: *presentation and write a data management plan exercise*.
- (4) Two-day train the trainers event on *Looking after and managing your research data*.
- Overview of the UK Data Archive's managing and sharing training materials: *presentation*.
 - Practical data management planning and funder policy context: *presentation and exercises*.
 - Consent and ethics: *presentation and group exercises*.
 - Data anonymisation: *presentation and group exercises*.
 - Formatting data: *presentation, formats quiz, file naming and transcription exercises*.
 - Documenting and contextualising your data: *presentation and context exercise*.
 - Data copyright: *presentation and research scenario discussions*.
 - Storing and transferring data and data security: *presentation, software demonstrations and quiz*.
 - Training clinics: *discuss your personal data needs with the trainers*.

In these sessions, we observe that group work is important and can greatly aid the learning experience.

To incorporate such data management training into taught courses that may be constrained in terms of time or curriculum space, the modules fit alongside new and existing training provision. At undergraduate level, the typical 10-week course on quantitative methods likely does not have enough time to include data skills that sit outside the confines of data collection and analysis. For qualitative methods courses, a ten-week course has barely enough time to cover the practicalities of collecting, transcribing, analysing and reporting research data. At masters and doctoral level, it may become feasible to include half-day or one-day data skills training. As we suggested above, flexibility may be the key, for example adding appealing workshop activities to mainstream courses.

One of the most positive outcomes of teaching these materials has been the positive reception of qualitative researchers to distinct topics that might otherwise be deemed out of their comfort zone. At the same time, using real life examples of research that are relevant and topical can aid qualitative researchers to fully engage with some of the more technical topics. For areas like ethics and consent or

interview transcription, qualitative researchers may feel more confident engaging, perhaps even by reflecting on and challenging their own everyday practices (like only seeking consent for sole access to data, thereby precluding colleagues and research students from using their data in the future). However, topics like data security, encryption in transferring disclosive data between colleagues, and creating metadata can appear technically difficult. By engaging participants in practical and relevant activities we can demystify the complexity of some of the more technical digital data practices or perceived heavyweight topics, supporting learners in gaining these essential practical skills.

Who is best placed to teach?

Questions may be raised over whether methods teachers or research support staff are best placed to teach these materials, or whether learners should attend face-to-face training or self-study. Encouragingly, we have seen interesting examples of academics teaching research methods being creative in their classes by incorporating data reuse assignments that enable critical evaluation of secondary data and exposure to other researchers' data (Haynes & Jones, 2012; Kelly, 2012; Smith, 2008; Turton, 2012). Moreover, emerging literature on re-use of qualitative data, such as in a 2012 Special Issue of this journal on working with archived textual and visual material in social research helps to inspire students about the possibilities of confronting data they have not collected themselves (Crow & Edwards, 2012; Kynaston, 2005; Corti, Witzel, & Bishop, 2005).

Alternatively it may be that it is not the responsibility of methods teachers to cover these skills, but that of professional skills trainers. Many universities have set up skills-oriented courses to help amplify their students' employment prospects and to support lifelong learning. For example, SCUNL (2014), as an academic library-based enterprise, aims to support both research excellence and students' academic success and employability. They cover important areas of information and digital literacy. In the past year the topic of data management has been added to their list of priority areas. Indeed there is no reason why data management skills could not be run as a parallel track to any kind of employability or professional development-oriented training.

Finally, there is also the role of institutional or faculty research offices that are already charged with offering support to researchers on funding opportunities, grant applications and research ethics submissions; as well as providing training on academic skills such as writing grant applications and academic writing. As their portfolio of support has increased to provide active advice for writing data management plans, training on some data skills could be handled by them. Skills could include writing data management planning, costing for data management in grant applications, and ethical issues in collecting and reusing data. In similar ways, information systems service providers in universities can provide expertise and training in the more technical aspects of data management and sharing, such as good practices in data storage, backup, data encryption, data security measures and data transferring.

One of the greatest challenges of teaching on these topics is ensuring the link with current research practice (and therefore trainers' familiarity with research practices). Success may depend on the experience of the trainers and on their ability to speak the language of research of their audience; being confident to answer questions about ethics, fieldwork procedures, data validation and so on. Ultimately, any

of these actors are capable of teaching data management skills to students and researchers. The availability of teaching materials and exercises such as those presented by the authors enables many professionals to take on this form of teaching, either as part of or complementing the research methods curriculum.

Conclusion

Researchers are expected to adapt to new ways of working with data that require skills to accompany their existing analytics techniques. Data management skills can be seen as the building blocks for data science and in turn, a core set of skills for social scientists rather than the domain of dedicated data managers. All empirical researchers are data managers to some degree. As data sharing requirements come into play, many placing the research institution at the heart of responsibility for data sharing (EPSRC, 2011), the pressure to build capacity amongst staff and researchers across all research disciplines in data management is increasingly crucial. Trained teaching, research, and support staff can engage students and learners in data management, beginning a cycle of upskilling. Training provision for managing and sharing data is emerging at a fairly rapid pace, led by institutions charged with overseeing the sharing of data by researchers, in turn driven by conditions of funding for their research projects. However, a cultural change is still needed to encourage research methods teachers and trainers to embrace and capitalise on this emerging field of pedagogy.

In this paper we set out to demonstrate teaching strategies that can help enhance social scientists' data management skills. Higher-level needs of gaining proficiency and confidence in digital data handling can be met by translating them into practical guidance and training activities that embed into research (see also Silver & Woolf, 2015). With the uptake of these strategies in the UK and beyond, this approach is contributing to jump-starting the research methods curriculum for data management. At its heart, this pedagogy relies on easy-to-digest information on considered topics, coupled with grounded activities that reflect either the substantive or methodological interests of learners. Where possible, such modular activities complement existing research methods training approaches and instructional literature. Be it the ethical aspects of data sharing, collecting and entering data using software packages, or the importance of data documentation, all parts of research design and analysis have short and longer-term data management implications.

Finally, with pressure for more transparency in research, for increased reuse of existing data, and for handling confidential data well, the pedagogical approach outlined in this paper offer skills that can help researchers move more readily and confidently into this space. These transferable skills help researchers to be better organised in dealing with digital data sources, and be better prepared to work within ethical and legal frameworks that govern research data. Further, they offer a strategic contribution to the UK's research capacity building programme in the social sciences.

Acknowledgements

We would like to acknowledge our colleagues who have undertaken training in areas of research data management at the UK Data Archive, in particular, Libby Bishop who contributed to our training modules and the 2014 handbook on Managing and Sharing Data.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Louise Corti is an associate director of the UK Data Archive and Head of the Collections Development and Producer Support teams. She has directed ESRC research data policy and sharing activities since 1995 also leads on qualitative data activities at the Archive.

Veerle Van den Eynden manages the Research Data Management team at the UK Data Archive. She leads on a variety of research and development projects on research data management and actively supports ESRC data policy activities.

References

- Bishop, L. (2012). Using archived qualitative data for teaching: Practical and ethical considerations. *International Journal of Social Research Methodology*, 15, 341–350. doi:10.1080/13645579.2012.688335
- Blomfield, M. (2012). Ethics in economics: Lessons from human subjects research. *Erasmus Journal for Philosophy and Economics*, 5, 22–44. Retrieved from <http://ejpe.org/pdf/5-1-art-2.pdf>
- British Academy. (2012). *A position statement – Society counts, quantitative skills of social sciences and humanities*. London: Author.
- British Academy. (2014). *High level strategy group for quantitative skills*. Author [Website]. Retrieved from http://www.britac.ac.uk/policy/High_Level_Strategy_Group_for_QS.cfm
- Buckley, J., Brown, M., Thomson, S., Olsen, W., & Carter, J. (2015). Embedding quantitative skills into the social science curriculum: Case studies from Manchester. *International Journal of Social Research Methodology*, 18(5), 495–510.
- Centre for Multilevel Modelling. (2014). *Learning environment for multilevel methodology and applications*. Retrieved from <http://www.bristol.ac.uk/cmm/learning/online-course/>
- Corti, L., Van den Eynden, V., Bishop, L., & Woollard, M. (2014). *Managing and sharing research data: A guide to good practice*. London: Sage. ISBN 978-1-44626-726-4.
- Corti, L., & Watkins, W. (2004, Summer/Fall 4-5). Guest editors' notes. In L. Corti & W. Watkins (Eds.), *IASSIST quarterly* [Special edition on statistical literacy]. Retrieved from http://www.iassistdata.org/downloads/iqvol282_3editor_0.pdf
- Corti, L., Witzel, A., & Bishop, L. (Eds.). (2005, special issue). Secondary analysis of qualitative data. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, 6. Retrieved from <http://www.qualitative-research.net/index.php/fqs/article/view/498>
- Crow, G., & Edwards, R. (Eds.). (2012, special issue). Perspectives on working with archived textual and visual material in social research. *International Journal of Social Research Methodology*, 15. doi:10.1080/13645579.2012.688308
- Digital Preservation Coalition. (2014). *Digital preservation training programme*. Author [Website]. Retrieved from <http://www.dpconline.org/training/digital-preservation-training-programme>
- DCC. (2013). *Digital curation training for all*. Author [Website]. Retrieved from <http://www.dcc.ac.uk/training>
- EPSRC. (2011). *EPSRC policy framework on research data*. Author [Website]. Retrieved from <http://epsrc.ac.uk/about/standards/researchdata>
- ESRC. (2012). *Secondary data analysis initiative*. Author [Website]. Retrieved from <http://www.esrc.ac.uk/research/skills-training-development/sdai/>
- ESRC. (2013). *Big data network*. Author [Website]. Retrieved from <http://www.esrc.ac.uk/research/major-investments/Big-Data/>
- ESRC, BSA and HAPS. (2010). *International benchmarking review of UK sociology*. Swindon: ESRC.
- Green, A., & Gutman, M. (2006). Building partnerships among social science researchers, institution-based repositories and domain specific data archives. *OCLC Systems and Services: International Digital Library Perspectives*, 23, 35–53. Retrieved from <http://deepblue.lib.umich.edu/handle/2027.42/41214>
- Hamilton, P., Gossman, P., & Southern, K. (2014). *Developing innovative support structures for students undertaking small-scale research projects in work settings*. The Higher

- Education Authority [Website]. Retrieved from https://www.heacademy.ac.uk/sites/default/files/resources/Glyndwr_final.pdf
- Haynes, J., & Jones, D. (2012). A tale of two analyses: The use of archived qualitative data. *Sociological Research Online*, 17(2), 1. Retrieved from <http://www.socresonline.org.uk/17/2/1.html>
- Inter-University Consortium for Political and Social Research. (2009). *Age and attitudes about the rights of homosexuals: A data-driven learning guide*. Author, University of Michigan [Website]. Retrieved from <http://dx.doi.org/10.3886/gayrights>
- Kelly, A. (2012). *Introducing sociology students to quantitative research methods*. UK Data Service Case Study, University of Essex [Website]. Retrieved from <http://ukdataservice.ac.uk/use-data/data-in-use/case-study/?id=19>
- Kilburn, D., Nind, M., & Wiles, R. (2014a). Learning as researchers and teachers: The development of a pedagogical culture for social science research methods? *British Journal of Educational Studies*, 62, 191–207. doi:10.1080/00071005.2014.918576
- Kilburn, D., Nind, M., & Wiles, R. (2014b). *Short courses in advanced research methods: Challenges and opportunities for teaching and learning*. Southampton: National Centre for Research Methods Report, NCRM, University of Southampton. Retrieved from http://eprints.ncrm.ac.uk/3601/2/advanced_methods_short_courses.pdf
- Kirton, A., Campbell, P., & Hardwick, L. (2014). *Developing applied research skills through collaboration in extra-academic contexts*. The Higher Education Authority [Website]. Retrieved from <https://www.heacademy.ac.uk/sites/default/files/resources/Liverpool.pdf>
- Kottmann, A. (2011). Reform of doctoral training in Europe: A silent revolution? In J. Enders, H. de Boer, & D. Westerheijden (Eds.), *Reform of higher education in Europe* (pp. 29–43). Rotterdam: Sense.
- Kynaston, D. (2005, January). The uses of sociology for real-time history. *Forum: Qualitative Social Research*, 6(1). Retrieved from <http://www.qualitative-research.net/index.php/fqs/article/view/503>
- Leston-Bandera, C. (2013). Ten tips to develop engaging undergraduate research methods teaching. Blog post. *Political Insight*. Political Studies Association [Website]. Retrieved from <http://www.psa.ac.uk/political-insight/blog/ten-tips-develop-engaging-undergraduate-research-methods-teaching>
- Library of Congress. (2013). *Digital preservation outreach and education*. Library of Congress [Website]. Retrieved from <http://www.digitalpreservation.gov/education/curriculum.html>
- MacInnes, J. (2012). Quantitative methods teaching in UK higher education: The state of the field and how it might be improved. *HEA Social Sciences teaching and learning summit: Teaching research methods*. Retrieved from http://www-new1.heacademy.ac.uk/assets/documents/Events/SS_assets/Blog/MacInnes_fullpaper1.pdf
- National Centre for Research Methods. (2014). *Current research programme*. University of Southampton [Website]. Retrieved from <http://www.ncrm.ac.uk/research/>
- Nuffield Foundation. (2012). *Programme backgrounds – Promoting a step change in the quantitative skills of social science undergraduates*. London: Nuffield Foundation.
- Nuffield Foundation. (2013). *Why Q-step? Hopes for wider change*. Presentation. Author. Retrieved from https://www.heacademy.ac.uk/sites/default/files/resources/4_4_Nuffield_panel.pdf
- Research Councils UK. (2011). *Common principles on data policy*. Author. Retrieved from <http://www.rcuk.ac.uk/research/datapolicy/>
- Research Councils UK. (2012). *Research councils UK policy on open access*. Author. Retrieved from <http://www.rcuk.ac.uk/research/openaccess/policy/>
- Royal Statistical Society. (2014). *Getstats – Statistical literacy for all*. Author. Retrieved from <http://www.getstats.org.uk>
- SCONUL. (2014). *Shared and collaborative services strategy group*. Author [Website]. Retrieved from <http://www.sconul.ac.uk/page/shared-and-collaborative-services-strategy-group>
- Silver, C., & Woolf, N. (2015). From guided-instruction to facilitation of learning: The development of five-level QDA as a CAQDAS pedagogy that explicates the practices of expert users. *International Journal of Social Research Methodology*, 18, 527–543.
- Sloan, L. (2013). *Innovation in the assessment of social science research methods, social sciences blog*. The Higher Education Academy. Retrieved from <http://blogs.heacademy>

- ac.uk/social-sciences/2013/11/11/innovation-in-the-assessment-of-social-science-research-methods/
- Smith, E. (2008). *Using secondary data in educational and social research*. Oxford: OUP Press.
- Strayhorn, T. (2009). The (in)effectiveness of various approaches to teaching research methods. In M. Garner, C. Wagner, & B. Kawulich (Eds.), *Teaching research methods in the social sciences* (pp. 119–130). Farnham: Ashgate Publishing Group.
- Trzesniewski, K., Donnellamn, B., & Lucas, R. (Eds.). (2011). *Secondary data analysis: An introduction for psychologists*. Washington, DC: APA.
- Turton, J. (2012). *Getting sociology students into archived qualitative data*. UK Data Service Case Study, University of Essex. Retrieved from <http://ukdataservice.ac.uk/use-data/data-in-use/case-study/?id=107>
- UK Data Service. (2014a). *Case studies of re-use*. Author [Website]. Retrieved from <http://ukdataservice.ac.uk/use-data/data-in-use.aspx>
- UK Data Service. (2014b). *Prepare and manage data*. Author [Website]. Retrieved from <http://ukdataservice.ac.uk/manage-data>
- UK Data Service. (2014c). *Events at the UK Data Service*. Author [Website]. Retrieved from <http://ukdataservice.ac.uk/news-and-events/events.aspx>
- University of Bristol. (2014). *Data.bris*. Author [Website]. Retrieved from <http://data.bris.ac.uk/research/>
- University of Edinburgh. (2014). *Research data management guidance*. Author [Website]. Retrieved from <http://www.ed.ac.uk/schools-departments/information-services/research-support/data-library/research-data-mgmt>
- University of Michigan. (2014). *Social network analysis course*. Coursera [Website]. Retrieved from <https://www.coursera.org/course/sna>
- University of Southampton. (2014). *Web science: How the web is changing the world course*. Author, FutureLearn [Website]. Retrieved from <https://www.futurelearn.com/courses/web-science-2014-q3>
- Urban Big Data Centre. (2014). *Overview*. Author [Website]. Retrieved from <http://ubdc.ac.uk/>
- Vartanian, T. (2010). *Secondary data analysis*. New York, NY: OUP.
- Vitae. (2013). *Transforming professional development for researchers: Vitae achievements and impact 2008–2012*. Author [Website]. Retrieved from <https://www.vitae.ac.uk/vitae-publications/reports/report-achievements-vitae-2008-2012.pdf>
- Williams, M., Payne, G., Hodgkinson, L., & Poole, D. (2008). Does British sociology count? Sociology students' attitudes toward quantitative methods. *Sociology*, 42, 1003–1021.