Big qualitative data: creative ways of working with qualitative analysis at scale

Authors: Sophie Woodward, Laura Radcliffe, Kirsty Pattrick, Sarah Bulloch, Mathew Gillings, Nicola Madge, Anna Tarrant, Ellie Turner-Kilburn, Emma Casey, Rosalind Edwards, Laura Fenton, Vicky Hill, Kahryn Hughes, Dawn Knight, Neve Meats, Hannah Musiyarira, David Oakey, Justyna Robinson, Jessica Scantlebury, Susie Scott, Christina Silver, Leighann Spencer, Rachel Thomson, Susie Weller, Emma Wincup, Natalie Wood.

What is this resource and who is it for?

This resource is for anyone who is committed to the importance of qualitative research but also recognises the complexities of working with 'qualitative data. In particular in this resource we pick up the challenges and possibilities of working with 'big qualitative data'. This resource is for a range of career stages: you might be doing a PhD, research in the public sector or be a well-established academic. The kinds of data you work with could include diary studies, creative projects using multiple methods or analysing secondary data in the Mass Observation Archive. There are multiple forms that big data takes when we think about qualitative research and in this resource we try to speak to this range. This document doesn't provide all the answers but gives you the challenges, debates and resources you can consult and steps to take. There are some contentious debates within the field that we do not seek to resolve but instead indicate what these concerns are, and some resources or things to think about. The resource is authored by the organising team (Sophie Woodward, Laura Radcliffe and Kirsty Pattrick) and those who attended an Innovation Forum at The Keep in September 2025, as well as those who were unable to attend but have commented on this document. We are authors with differing backgrounds, interests and approaches but we come together here to put those interests and expertise into dialogue. We invite you, the reader, to join in this dialogue.

How did this resource come about? What do we mean by 'big qualitative data'?

This resource came out of an Innovation Forum funded by NCRM; the starting point was discussions between the organising team (Sophie Woodward, Laura Radcliffe, Kirsty Pattrick and in the early stages Ali Hanbury) around the desire to articulate and explore qualitative approaches to what has been termed 'big data'. The need for an event centring qualitative approaches arose from an awareness of the

tendency in discussions of big data to presume that 'big data' means quantitative approaches. Quantitative methods have been usefully deployed to offer broad and generalised understandings of various social phenomena through the generation of large scale datasets. There are therefore evident ways in which quantitative approaches have been useful in the analysis of large datasets. However, with the increasing tendency and expectation for qualitative data to be archived, we are faced with many large and complex qualitative datasets. Some of these, such as the Mass Observations Archive have been around for a long time, and others are explicitly framed as being longitudinal datasets - such as the Timescapes archive. Timescapes was proactively built to house qualitative datasets that were being produced at the time of its inception and now offers a repository specifically for qualitative longitudinal datasets. The longer these archives/datasets accumulate data, the larger and more challenging they can become. In other cases, a desire to work across contexts through comparative approaches, or to compare datasets over time, leads to researchers working with large amounts of qualitative data.

In addition, utilising creative methods such as diaries, poetry, object-based interviews, and creative mapping, generates large amounts of varied forms of data which are textured and multi-modal, from which researchers can gain deep and rich insights and knowledge. These are often complex and generate multiple data sets, and qualitative researchers face a number of possibilities as well as challenges. They raise fundamental questions such as: what counts as big data? They also raise tricky questions around how we deal with multiple data forms. Epistemic approaches to the analysis of such data are complex. Analytical practices vary, and the development of skills, techniques, and approaches to this data is a growing field.

As researchers engage with increasingly large and multimodal qualitative datasets, new computational tools, including new forms of machine learning and generative AI, are being explored. Some see potential for these tools to augment searchability, organisation, researcher reflexivity, and/or visualisation of complex data, when used critically and with care as part of the analytical process. However, others raise epistemological and ethical concerns about environmental costs, automation, and the risk of the interpretive, reflective labour that defines qualitative work being eroded in favour of faster non-human analysis that may lead to analysis depth or significance (this debate is played in Jowsey et al 2025 and a response from Friese, 2025). This Innovation Forum event and subsequent resource acknowledges these active debates, recognising that approaches to big qualitative data will continue to evolve, and that critical, reflexive dialogue remains essential.

We chose to run an NCRM Innovation Forum as an opportunity to bring together researchers with varied approaches and data to discuss the challenges and possibilities of working with big qualitative data. We held an in-person event which was attended by around 35 invited researchers with specialisms in

different disciplines and methods where, through discussion and hands-on activity, we explored the questions including, what is big data, and what challenges and possibilities does working with this data afford? The day had the following objectives:

1. To facilitate knowledge exchange between experts in creative and qualitative research methodologies, focusing on the challenges and opportunities of "big qual" analysis.

2. To explore and compare diverse analytical approaches for handling large qualitative datasets generated through creative research methods and to increase capacity among attendees to apply innovative analytical approaches to their own "big qual" research projects.

'Big' Qualitative Data: Critical Questions for Researchers

Prior to the Innovation Forum event, and to support the in-person discussions on the day, we invited four qualitative researchers/research teams to create short preparatory videos introducing their own experiences of working with large or complex qualitative datasets. You can access each of these videos as part of this resource. During each video contributors reflected on:

- The types of "big" qualitative data they work with and the challenges these pose;

- Their analytical approaches and what these enable;

- The limitations of their approaches;

- The questions they would like others to consider; the unresolved conceptual, ethical, and methodological puzzles that warrant wider debate.

The final prompt - **"what questions do you want others to think about?"** - proved particularly generative. Many of the issues raised converged or spoke to similar concerns and critical questions that qualitative researchers are currently grappling with in the context of 'big' qualitative data. We collated these questions into five core questions to guide discussions on the day.

Together, these questions capture key tensions and possibilities in working with "big" qualitative data, spanning ethics, epistemology, technology, practicality and creative multimodality. The following sections introduce each question and summarise the discussions and key insights that emerged:

Question 1: How can we ensure qualitative data remains central to 'big data' conversations, and what unique value can big qual bring?
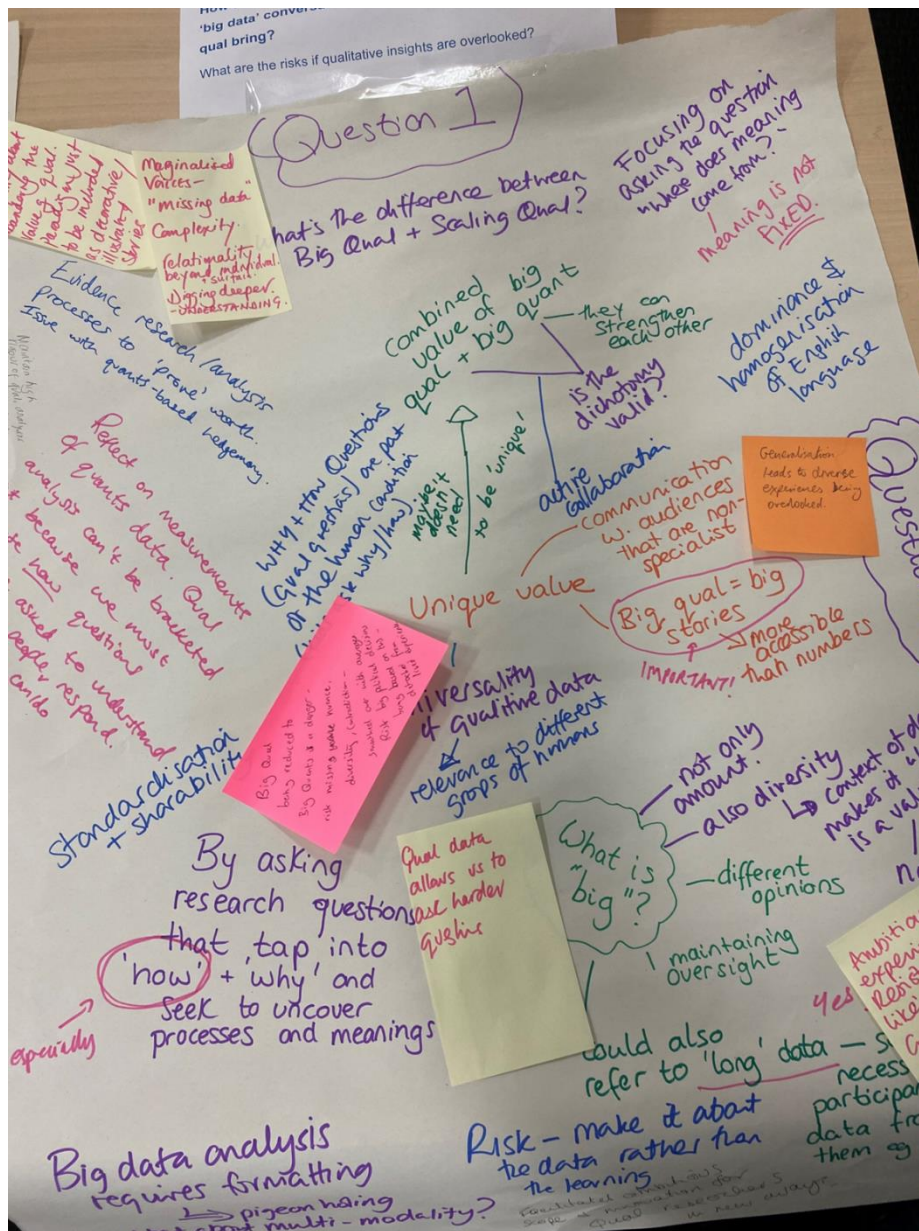

Image 1: Photo of discussions of question 1.

The discussions generated more questions:

- What's the difference between Big Qual and Scaling Qual?

- Is the dichotomy between qual and quant helpful or even valid?

- What is 'big'? (this is a core question we explored on the day as it can refer to: diversity of data, longitudinal engagements).

Through these discussions we articulated more closely what qualitative approaches to big data could offer, which included distinctive contributions such as:

- Focusing on asking the question "where does meaning come from?" – meaning is not fixed.

- By asking research questions that tap into the how and why, to uncover processes and meanings.

- The potential to foreground a diverse array of voices and perspectives, including those from more marginalised or under-represented communities.

- Qualitative approaches can often allow us to centre stories which can be a particularly powerful way to communicate to multiple different audiences who are not specialists. In turn this allowed us to redefine 'big' in a way that isn't just captured by numbers.

- The combined value of quantitative and qualitative methods and the importance of considering how they can complement one another.

- Qualitative approaches can include and draw on collaborative, artistic, or interdisciplinary processes.

What counts as data?

Questions around analysis also entailed discussions over what counts as evidence, and as data, and in turn how we might use this data. We discussed:

- The importance of standardisation and shareability for multiple different forms of data and some of the challenges this might pose (e.g. how standardisation limits possibilities for nuance/depth).

- How we use evidence/analyse in a way that retains nuance.

- Versatility of qualitative data which may allow relevance to a wide range of audiences (i.e. multiple forms of qualitative data).

Diversity:

We discussed diversity of data forms (e.g. interviews, observations, images, etc.) as well as diversity of voices and perspectives within the data.

- Big qualitative data can enable the inclusion of marginalised voices, including the ability to bring together multiple projects and contexts.

- 'Big' cannot just be reduced to the amount of data, but also the diversity of data/voices.

- Generalisation can lead to diverse experiences being overlooked.

- Dominance and homogenization of the English language which is manifest in multiple ways. One way is that if a big-qualitative infrastructure (for example platforms, coding software, datasets, search tools, AI-supported analysis) are built around English then we might risk researchers privileging English sources, voices, texts, and interactions.

Question 2: How can we ensure that 'big' qualitative data analysis is ethical and contextually grounded, retaining its qualitative integrity?

Image 2: Photo of discussions of question 2.

Discussions surrounding question two fell into five broad, overlapping areas as outlined below:

1. Maintaining epistemological and methodological integrity

- Ethical and analytical decisions must align with the philosophical underpinnings of the research.

- 'Big qual' analysis should not default to quantitative logic or generalisability but remain grounded in interpretivist commitments to depth, multiplicity, and meaning.

- Recognise that what counts as 'big' or 'context' varies across communities and disciplines - ethics should always be situated.

2. Maintaining contextual depth and individual voices

- Prioritise stories rather than fragments; preserve the richness and materiality of data.

- Capture affective and embodied dimensions of qualitative material, not just textual content.

- Remain attentive to minority and marginalised voices - resist the pull towards consensus or thematic aggregation.

- Vignettes or 'full pictures' of individuals can help retain narrative integrity while protecting anonymity.

3. Ensuring transparency and reflexivity

- Document analytical decisions and reflexive notes throughout - transparency of process.

- Keep fieldnotes rich and inclusive (capture everything, not only what seems relevant at the time).

- Build in 'slow time': robust qualitative analysis takes time and cannot be rushed to fit 'big data' expectations.

4. Considering ethics of re-use and AI-assisted analysis

- Treat re-use of qualitative data with care: don't assume consent covers future uses or AI modelling.

- Ethical review processes may need to evolve to accommodate retrospective consent or secondary analysis.

- If using AI tools, critically consider *how* outputs are produced and what interpretive labour or nuance could be lost. Consider carefully *how* AI is used and the critical interpretive role of the researcher.

- Ask: what can only humans do in the interpretive process?

5. Reflections on inclusion, and the scope and value of 'big qual' data, to keep in mind during data analysis

- Large qualitative datasets offer opportunities for inclusion - bringing in more diverse, everyday lives.

- Reflect on the purpose: not to generalise, but to broaden understanding and multiplicity of voices.

- Ethics is relational - it depends on the community, the dataset, and the researchers' interpretive commitments.

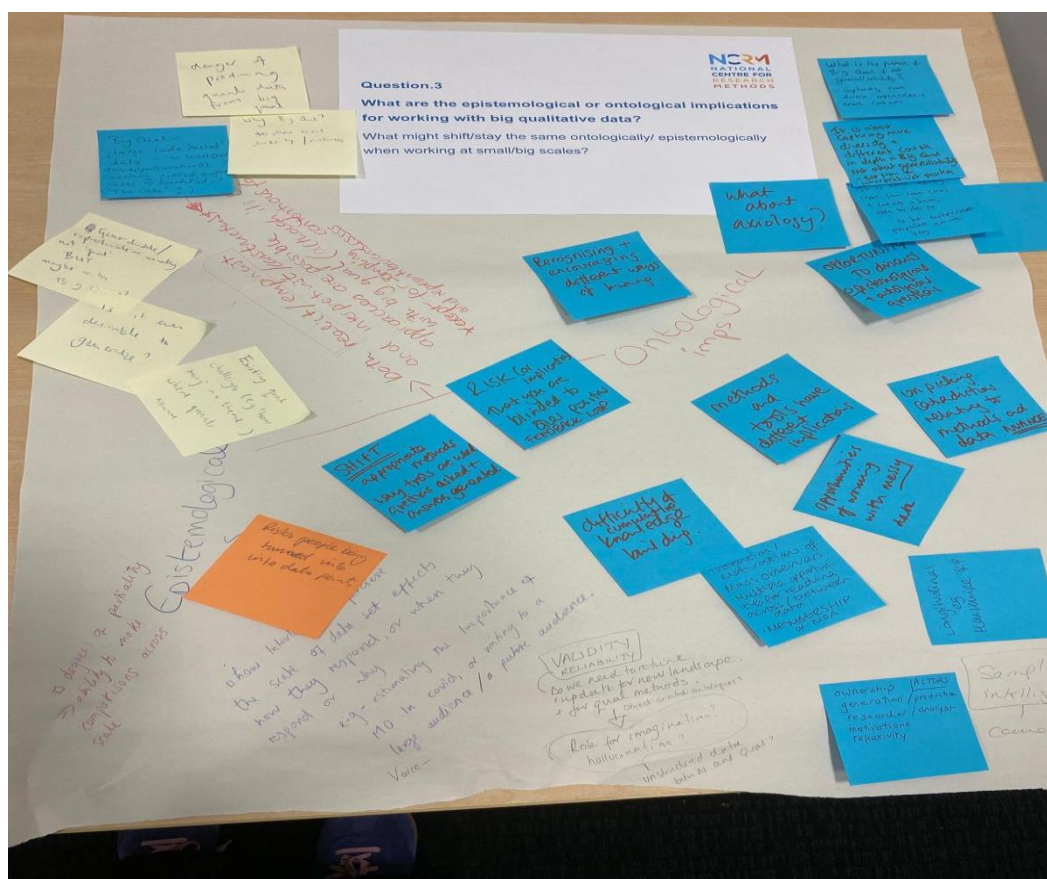# Question 3: What are the epistemological and ontological implications for working with big qualitative data?

Image 3: Photo of discussions of question 3.

Discussion looked at the opportunities big qualitative data offers with its richness; unpicking contradictions relating to methods and data, whilst encouraging and recognising different ways of knowing:

- Big qual should not be about generalisability, especially from an interpretivist position. Usually qualitative data is not aiming to be generalisable - but is this true for big qual data? Yet, is it even desirable to generalise?

- Instead should the aim be to capture more diversity and different contexts in-depth; the nuances, whilst offering the ability to make comparisons across scale?

- Both realist/empiricist and interpretist/constructivist approaches are possible with big qual.

- Axiology was suggested as another philosophical approach - a focus on values and what we consider good, worthwhile, or valuable research.

Challenges

- Risk (or implicated risk) arose as a challenge, where one becomes blinded to another position, leading to a 'Feedback Loop'.

- There can be a danger of producing a quantitative data set from big qualitative material, which risks people being turned into information data points.

- Difficulty of cumulative knowledge building.

- Methods and tools have different implications. It requires a shift to appropriate methods and the way tools are used, the questions asked and answers generated. Points were made on:

  - the role of participants and their intention/motivations to contribute

  - ownership of the data; generation/production, researcher/analyst, motivations, reflexivity.

When exploring appropriateness and rigor of approaches/methods, the discussion led to a bigger question of whether we need to re-think and update positions for a new landscape and for qualitative methods such as object orientated ontologies.

Question 4: (How) Might computational support make the analysis of big qualitative diary data more sustainable, while retaining individual 'momentary' narrative context and *capturing change over time*?

For many researchers, the relation between the researcher and any computational tool, including AI, is central to their practices. This relation should always be reflexive. Computational tools may be able to assist organisation or exploration but should never substitute the interpretive and reflexive dimensions of qualitative inquiry. Discussions indicated that where used, programmes should assist, not replace and AI should support, not substitute human researchers as we are finding ways of working together. We need to know what questions we are asking of the data, and how this translates to questions AI can respond to and work with. This means that if the researchers themselves are using AI for their own analysis they will build an awareness of the fieldwork and

other contextual information when asking questions of AI. If this is for secondary data we need to make sure we build in the demographic and contextual information into the meta data to make this possible. We need to make sure there is knowledge of what the research questions were in the broader research context.

Image 4: Photo of discussions of question 4.

Challenges of:

- semantic drift (shifts in meanings of words) and the flattening of meanings when data is compressed. AI can result in both of these happening.

- lack of context in use of many computer programs and therefore loss of nuance.

Some challenges are also possibilities:

- Computer programs can offer speed but what is lost with speedy analysis? What about quality? What about the advantages of slowness? This balance between efficiency and interpretive integrity is at the heart of current debates on the role of AI in qualitative research.

Ethics and accountability:

- If AI is involved, who did the analysis/authoring? We must show accountability by being explicit about the use of AI.

- Sustainability in terms of environmental impact. How is this taken into account? How can we support the use of AI that is thoughtful and not excessive?

We discussed some more specific and technical issues:

- Digitisation of data, data management and longitudinal tracking over time. If we are thinking about support with analysis then this aspect of data needs to be considered. How can we use meta data to help in our analysis?

- How might AI help us with moving between big/small, macro/micro? How we move between scales needs rethinking with different tools.

- Software tools might help speed up creation of 'trajectory diagrams' (Spencer et al, 2021) after researchers' own interpretive coding and rough plotting.

- Could existing software (e.g. Sketch Engine) be used alongside researcher interpretation to help give *initial* points of analysis to compare over time?

- CAQDAS tools may support integration and comparison, augmenting analysis rather than replacing it.

- Capturing and retaining thematic change over time is crucial to visualise trajectories and meanings.

- We also discussed ways in which we could develop productively in this area and we suggested that there is the need for training in how to use computer programmes/AI, and also collaborations between social sciences and data sciences to broaden understanding across disciplines and

learn from one another. We also suggested that the programs themselves should be designed and workable for people who are working with a diverse range of qualitative data.

Question 5: How can we analyse big qualitative datasets that include diverse creative contributions (e.g. photographs, artwork, videos, narratives)?
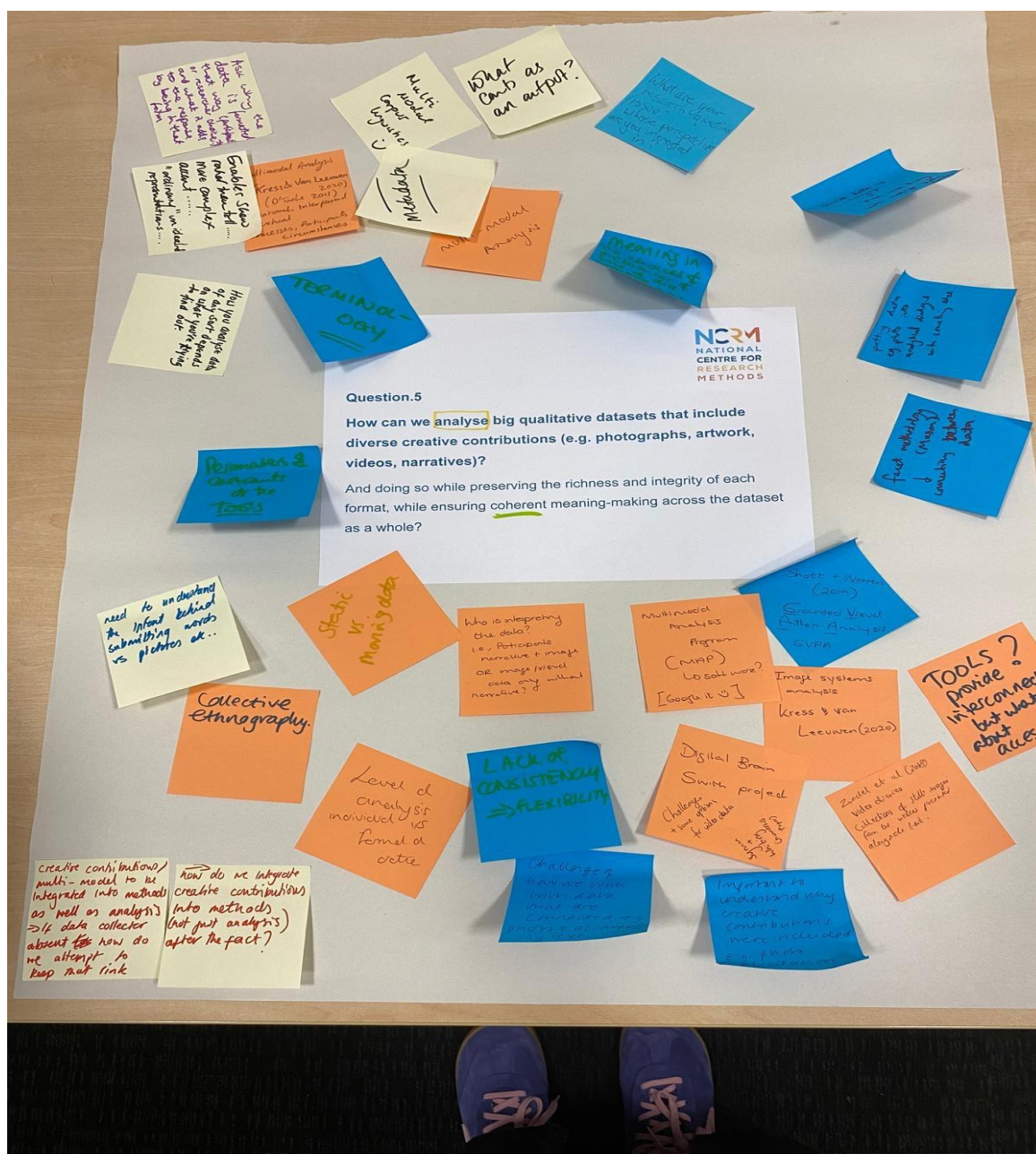

Image 5: Photo of discussions of question 5.

This discussion led to further questions on context. For example, looking at the need to understand the intent behind submitting textual or visual data to the Mass Observation Archive. To explore analysis it's important to understand why creative submissions were included in the first place (e.g. photographs or drawings). It's important to understand the meaning behind the choice of creative medium.

When approaching the data:

- What are your research questions? Whose perspective are you interested in? How you analyse data of any sorts depends on what you're trying to find out.

- Who is interpreting the data and what is being interpreted? (i.e. participant's narrative and image, or image/visual data only, without narrative).

- Ask why the data is formatted that way (participant or researcher choice?) and what it adds to the response by being in that form.

There was talk around creative contributions and multi-modal approaches being integrated into methods as well as analysis.

Questions arose with regards to secondary data:

- If the data collector is absent, how do we attempt to keep that link?

- When working with diverse creative contributions, what counts as an output?

Challenges

Regarding the data, this can lead to:

- the lack of consistency where there is flexibility in formats.

- variety in terminology when working with different forms of data.

When it comes to the analysis:

- how to manage the parameters and constraints of the tools.

- how to trim data that are connected (e.g. photo and accompanying text).

- Tools provide interconnectivity but what about accessibility?

Analysis

This generated a fruitful discussion on the different types of analysis that could be used:

- Initially you could start by reading it all, taking a random sample or shifting the order in which you read.

- Sampling using biographical metadata.

- Look at the different formats separately; for example, photos then illustrations, using image analysis.

- Looking at the meanings in the choices and preferences of different creative contributions.

- Multi-modal approach, combinations to provide deeper insights.

- Corpus linguistics.

- Key concept approach (Robinson et al, 2023).

- Collective ethnography.

- Using case studies with multiple data forms.

- Putting data (e.g. photo) into analytical dialogue with something else.

- Facet methodology (Mason, 2011) connecting between data.

- Static vs moving data.

- Ideational, Interpersonal, Textual.

- Process, participants, circumstances.

- It enables a 'show' rather than 'tell' with a more complex account - 'Showing' ordinary unideal representations.

Working With Data: Mass Observation as a case study

The Innovation Forum took place at the Mass Observations Archive and one of the workshop tasks involved using materials from the archive to explore methods of approaching this type of data. Samples were selected that incorporated narratives with illustrations and photographs.

Mass Observation maintains a national panel of self-selecting volunteer writers (known as 'Mass Observers'). There are approximately 600 who respond to open questionnaires ('directives') on a wide range of personal, social and political topics (anonymously, with a unique individual code). For this exercise there were samples of data from the Summer 2025 Directive. This directive contained three topics:

- An object; a timepiece.

- A body part; hand.

- A place; kitchen.

The workshop participants were given 3 individual responses to one of the above topics. This provided a mix of data, in content, format and length, to aid discussion.

There was discussion on the richness of data; the flexibility of the 'directive' method which enables contributors to respond in a way that feels comfortable to them, in their own space and time; and the opportunity to interpret the questions and contribute in their own way. This generates an 'openness' to share and confide. It is a form of life writing, reflecting back on lived experiences, with thoughts and opinions of the 'now'.

The challenges that then arise with this method include the length of narrative, variety of format (inserting photographs within text, timelines, bullet points) and style (poetic, illustrative), mix of narrative with illustration and/or photograph, along with the size of the dataset (circa. 180 individual submissions).
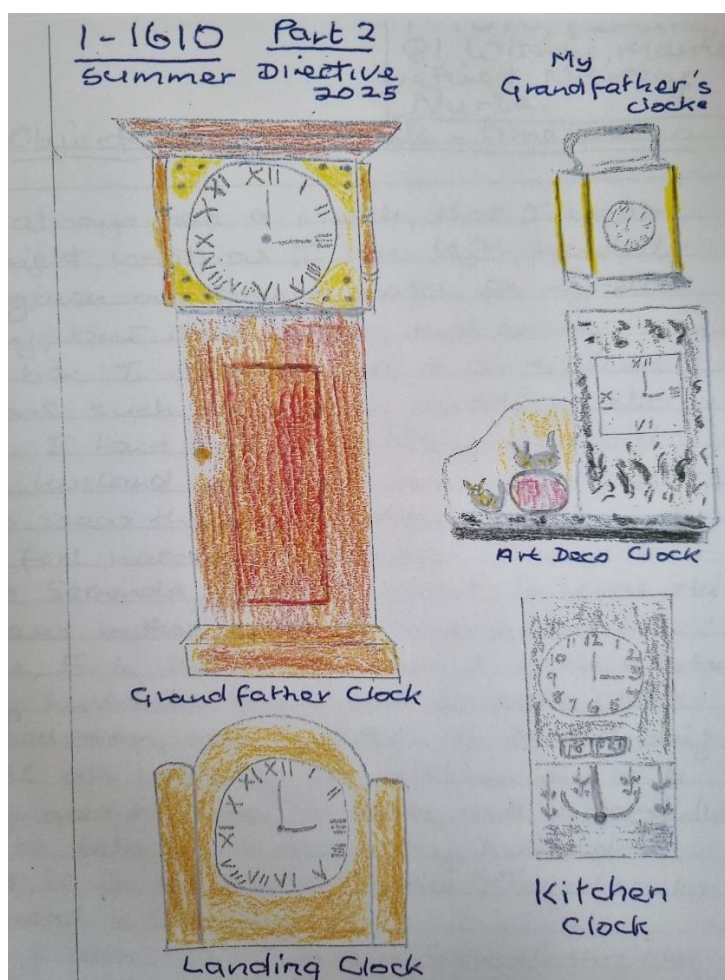
Image 6: Response from Mass Observer I-1610 to Summer Directive 2025 Part.2

* The scanned images are extracts from individual responses

Top 10 things my HANDS did for me today.

① Peeled an Orange.
② Held the dog lead on an early morning walk in the forest.
③ Turned a wrong-way-up beetle the right way up.
④ Planted radish seeds in my vegetable box, pushing them into damp, warm soil.
⑤ Rummaged in a drawer looking for stamps and found the long-lost sewing kit.
⑥ Sewed a button on my shirt.
⑦ Stroked the dog.
⑧ Crossed off more than half the things listed on my TO DO List.
⑨ Mixed the salad.
⑩ Turned the pages of the first chapter of a new book.

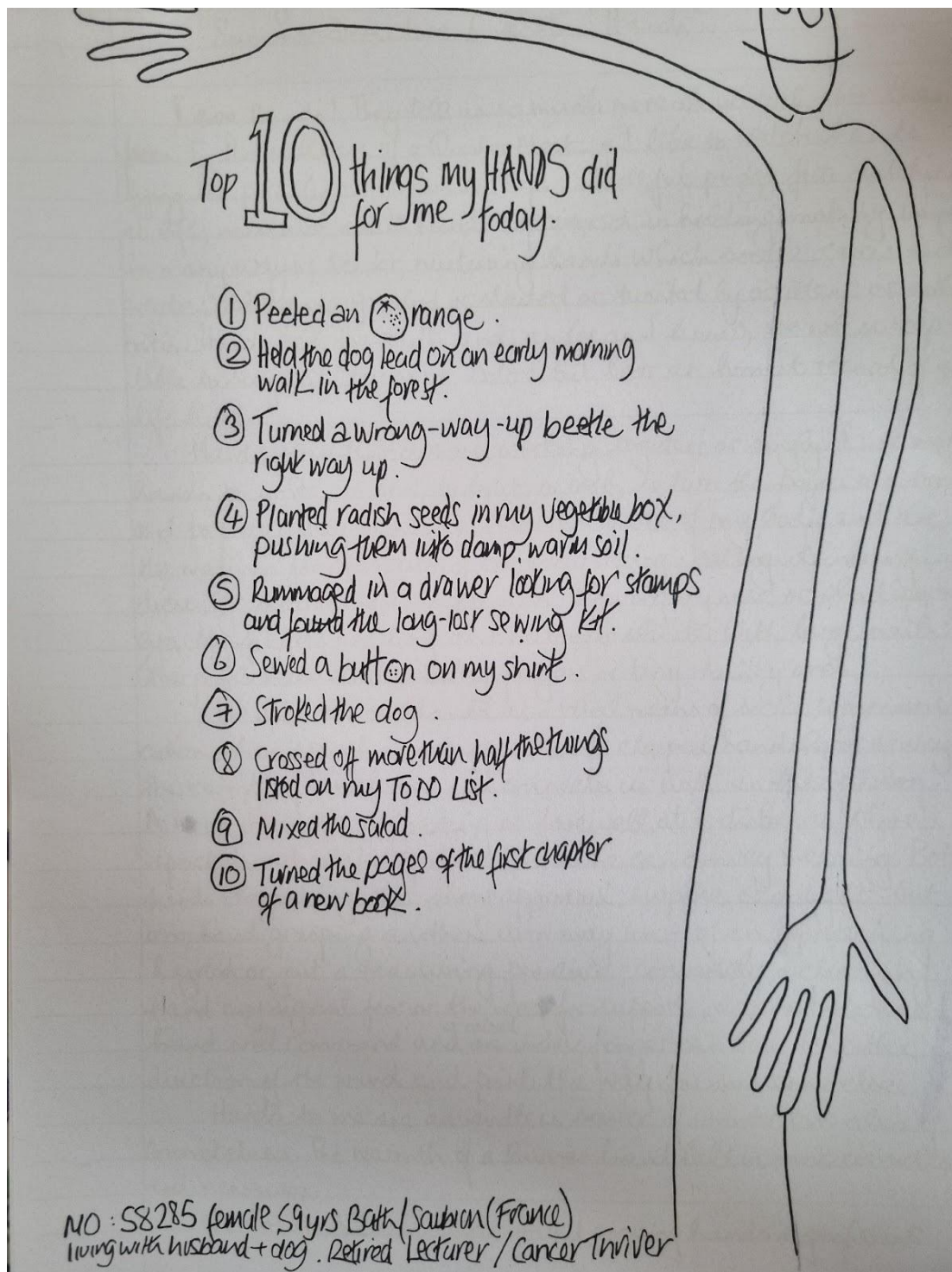NO: S8285 female 59yrs Bath/Saubion (France) living with husband + dog. Retired Lecturer / Cancer Thriver

Image 7: Response from Mass Observer S8285 to Summer Directive 2025 Part.1

Part 3.

W8217, cis-female, 57, Whiteley (Hants), lives with husband, university lecturer

Part 3: A Place in Your Life – a kitchen

Look

Over the last twenty-one years, I have spent lots of time in my kitchen. It is the heart of our home. It's a gathering place where the family come together. The place where we sit and eat together, celebrate special occasions and enjoy our favourite family dishes. For a while during the pandemic, it also became my office.

We had it all replaced in 2007 when I was starting my new role as a local authority office. The first week of my new job was probably not the best time to have a new kitchen fitted! I still think of it as a new kitchen but, in fact, it is almost twenty years old!

Our most recent kitchen facelift involved having paint mixed to match our Orla Kiely curtains. I love Orla Kiely. The designs remind me of my childhood and 70s designs.

Image 8: Response from Mass Observer W8217 to Summer Directive 2025 Part.3

Concluding comments

The NCRM Innovation Forum generated lively discussion and debate. The value of qualitative approaches to 'big data' was recognised, as was the need for further development of methodologies in this area. The workshops addressed the discussion topics with creativity, raising additional questions and concerns in the process. This resource presents the points drawn from the discussions, often verbatim, and provides the basis for further exploration of the field.

There are many different approaches to qualitative analysis of 'big data' as the day made clear. This was also exemplified by the examples provided by the four videos that form part of this resource. Thus Justyna Robinson outlined the use of computer software to identify keywords and patterns, and Susie Weller and colleagues described their iterative process examining both breadth and depth of the data, utilising quantitative methods and bringing together diverse datasets. Both of these approaches were appropriate when addressing a specific research question or theoretical stance. By contrast, Laura Radcliffe and Leighann Spencer, who described their diary data, and Kirsty Pattrick, who outlined the collection and analysis of diary and directive data from the Mass Observation project, placed greater focus on individual narratives, including visual items within datasets.

The reflections gathered here demonstrate how big qualitative data in its various forms continues to build momentum, enriched by multiple disciplinary, methodological, and epistemological perspectives. Such data has the capability to offer critical insights into understanding diverse social worlds, keeping lived experience and contextual meaning at the centre, as a vital complement to 'big' quantitative approaches. Sustaining this momentum depends on continuing these conversations, together.

Some further reading:

There is a wider body of literature which concerns some of the issues we have discussed in this article and how these relate specifically to Qualitative Secondary Analysis. On the Timescapes webpage, Kahyrn Hughes and Anna Tarrant have created a comprehensive reading lists that can be accessed here to think about these broader methodological debates:

Reading-list-for-QSA.pdf

Approaches to 'Big' Qualitative Data Analysis:

Davidson, E., Edwards, R., Jamieson, L., & Weller, S. (2019). Big data, qualitative style: a breadth-and-depth method for working with large amounts of secondary qualitative data. *Quality & Quantity*, *53*(1), 363-376.

Edwards, R., Davidson, E., Jamieson, L., & Weller, S. (2021). Theory and the breadth-and-depth method of analysing large amounts of qualitative data: A research note. *Quality & Quantity*, *55*(4), 1275-1280.

Gamieldien, Y., Case, J. M., & Katz, A. (2023). Advancing qualitative analysis: An exploration of the potential of generative ai and nlp in thematic coding. Available at SSRN 4487768.

Gillings, M., Mautner, G., and Baker, P. (2023). *Corpus-Assisted Discourse Studies.* Cambridge University Press.

Halford, S., Carr, L., Jenner, S., Aubin Le Quéré, M., & Randazzo, C. (2025, October 22). Qualitative analysis and large language models [Webinar]. *National Centre for Research Methods* (NCRM). https://www.ncrm.ac.uk/resources/video/?id=4995

Hamilton, L., Elliott, D., Quick, A., Smith, S., & Choplin, V. (2023). Exploring the use of AI in qualitative analysis: A comparative study of guaranteed income data. *International journal of qualitative methods,* 22, 16094069231201504.

Radcliffe, L., & Spencer, L. (2025). *Qualitative Diary Methods* (Qualitative Research Methods Series). SAGE Publications. [Includes 2 chapters on analysing qualitative diary data].

Robinson, J. A., Sandow, R. J., & Piazza, R. (2023). Introducing the key concept approach to the analysis of language: the case of REGULATION in Covid-19 diaries. *Frontiers in artificial intelligence*, *6*, 1176283.

Spencer, L., Radcliffe, L., Spence, R., & King, N. (2021). Thematic trajectory analysis: A temporal method for analysing dynamic qualitative data. *Journal of Occupational and Organizational Psychology*, *94*(3), 531-567.

Weller, S., Davidson, E., Edwards, R., & Jamieson, L. (2023). *Big qual: A guide to breadth-and-depth analysis*. Palgrave Macmillan.

Approaches to Multi-Modal Data Analysis:

Knoblauch, H., & Schnettler, B. (2012). Videography: Analysing video data as a 'focused' ethnographic and hermeneutical exercise. *Qualitative Research,* 12(3), 334-356.

Kress, G., & Van Leeuwen, T. (2020). *Reading images: The grammar of visual design*. Routledge.

Kress, G., & Van Leeuwen, T. (2020). The semiotic landscape. In *Language in use* (pp. 344-349). Routledge.

Mason, J., 2011. Facet methodology: The case for an inventive research orientation. *Methodological Innovations Online*, *6*(3), pp.75-92.

O'toole, M. (2011). *The language of displayed art*. Fairleigh Dickinson Univ Press. (2nd Edition).

Shortt, H. L., & Warren, S. K. (2019). *Grounded visual pattern analysis: Photographs in organizational field studies. Organizational Research Methods*, 22(2), 539–563.

Whiting, R., Roby, H., Symon, G., & Chamakiotis, P. (2018). 10 Participant-led video diaries. In: Bryman and Buchanan (Eds) *Unconventional Methodology in Organization and Management Research*, 190.Oxford, UK: Oxford University Press.

Woodward, S., 2016. Object interviews, material imaginings and 'unsettling' methods: Interdisciplinary approaches to understanding materials and material culture. *Qualitative Research*, *16*(4), pp.359-374.

Woodward, S., 2019. Material methods: Researching and thinking with things. London, Sage.

Ethical Considerations:

Edwards, R. (2025, November 24). *Qualitative research and generative AI: reflections on an NCRM seminar.* NCRM. https://www.ncrm.ac.uk/news/show.php?article=5869.

Friese, Susanne, Response to: "We Reject the Use of Generative Artificial Intelligence for Reflexive Qualitative Research" (November 01, 2025). Available at SSRN: https://ssrn.com/abstract=5690262 or http://dx.doi.org/10.2139/ssrn.5690262 .

Hesse, A., Glenna, L., Hinrichs, C., Chiles, R., & Sachs, C. (2019). Qualitative research ethics in the big data era. *American Behavioral Scientist*, *63*(5), 560-583.

Jowsey, T., Braun, V., Lupton, D., & Fine, M. (2025). We reject the use of generative artificial intelligence for reflexive qualitative research.(October 20, 2025).Available at SSRN: https://ssrn.com/abstract=5676462 or http://dx.doi.org/10.2139/ssrn.5676462.

Joyce, J. B., Douglass, T., Benwell, B., Rhys, C. S., Parry, R., Simmons, R., & Kerrison, A. (2023). Should we share qualitative data? Epistemological and practical insights from conversation analysis. *International Journal of Social Research Methodology*, *26*(6), 645-659.

Mills, K. A. (2019). *Big Data for Qualitative Research* (p. 90). Taylor & Francis..

Nguyen, D. C., & Welch, C. (2025). Generative Artificial Intelligence in Qualitative Data Analysis: Analyzing—Or Just Chatting?. *Organizational Research Methods,* 10944281251377154.

Whiting, R., Roby, H., Symon, G., & Chamakiotis, P. (2018). *Who's behind the lens? A reflexive analysis of roles in participatory video research*. *Organizational Research Methods*, 21(4), 766-793.

In addition to the above resources, the CAQDAS Networking Project Qual-AI pages includes a bibliography on the use of AI in qualitative data analysis, including ethical issues: [Qual AI] | [CAQDAS]

Potentially Helpful Computational Packages:

The following examples illustrate tools that some researchers use to support the management and exploration of large or multimodal qualitative datasets.

- Leximancer - https://www.leximancer.com/ - A text-mining tool that processes large bodies of textual data, identifies concepts via co-occurrence, and produces visual concept maps to show relationships between themes. Potentially, useful when you have extensive interview data (or similar), especially across multiple projects, and want an exploratory layer of insights, or an overview of that data, before deeper manual coding.

- MAXQDA - https://www.maxqda.com/ - A qualitative data analysis tool, supporting text, audio, and video data; coding, visualisation (code maps, word clouds), memoing, multi-user collaboration, and mixed-methods functionality.

- Sketch Engine - https://www.sketchengine.eu/ - Designed originally for corpus linguistics, this tool allows you to build or query large text corpora (billions of words), generate collocation/word-sketch analyses and compare text sets. The tool is potentially particularly relevant if your "big qual" data includes written textual material (e.g. diaries, large interview collections) and you want to trace language usage, change, or patterns across time or groups.

These are just some suggestions above, and there are more possibilities. The CAQDAS Networking Project provides information on a wide range of tools designed to support large or multimodal datasets.

Choosing a CAQDAS package | University of Surrey