OXFORD

Exploring the possibilities of Thomson's fourth paradigm transformation—The case for a multimodal approach to digital oral history?

Hannah K. Smyth[1], Julianne Nyhan[1,2]*, Andrew Flinn[1]

[1]UCL, Germany
[2]Technische Universität, Darmstadt, Germany
*Correspondence: Julianne Nyhan. E-mail: j.nyhan@ucl.ac.uk

Abstract

This article seeks to reorientate 'digital oral history' towards a new research paradigm, Multimodal Digital Oral History (MDOH), and in so doing it seeks to build upon Alistair Thomson's (Thomson, A., 2007, Four paradigm transformations in oral history. *Oral History Review*, 34(1): 49–70.) characterization of a 'dizzying digital revolution' and paradigmatic transformation in oral history (OH). Calling for a recalibration of the current dominance of the textual transcript, and for active engagement with the oral, aural, and sonic affordances of both retro-digitized and born digital OH (DOH) collections, we call for a re-orientation of the digital from passive to generative and self-reflexive in the human–machine study of spoken word recordings. First, we take stock of the field of DOH as it is currently conceived and the ways in which it has or has not answered calls for a return to the orality of the interview by digital means. Secondly, we address the predominant trend of working with transcriptions in digital analysis of spoken word recordings and the tools being used by oral historians. Thirdly, we ask about the emerging possibilities—tools and experimental methodologies—for sonic analysis of spoken word collections within and beyond OH, looking to intersections with digital humanities, sociolinguistics, and sound studies. Lastly, we consider ethical questions and practicalities concomitant with data-driven methods, analyses and technologies like AI for the study of sonic research artefacts, reflections that dovetail with digital hermeneutics and digital tool criticism and point towards a new MDOH departure, a sub-field that has potential to inform the many fields that seek patterns in audio, audio-visual, and post-textual materials, serially and at scale.

1 Introduction

A recent state of the field of digital history, which, like digital oral history (DOH), broadly maps to the digital humanities (DH; (Romein *et al.*, 2020)), outlines seven of its most widely used techniques. The authors are forthright that this is far from an exhaustive list, but it is telling that items are almost exclusively text oriented, whether in the type of sources or software used or the methods of data manipulation and analysis (Romein *et al.*, 2020). Likewise with another recent volume described by the editors as 'representative in that the used methodological research approaches correspond to the predominant directions of current digital history' (Paju *et al.*, 2020, p. 7), has focused almost exclusively on textual sources and methods.

Yet digitization is giving rise to an ever-increasing corpus of born digital and retrodigitized audio and audio-visual (AV) collections, and the resources in this

equation are, in turn, facilitating access to and interest in non-textual data sources and non-textual DH methods of analysis (Arnold *et al.*, 2021). Recent initiatives like the British Library's 'Unlocking Our Sound Heritage', part of a wider initiative to preserve UK sound heritage (British Library, 2019), are digitizing and making a range of sound recordings more accessible and potentially available for research. Yet, despite the 'sensory turn' across the academy extending into Galleries, Libraries, Archives, and Museum scholarship and practice (e.g. Howes, 2014), with its attention to the intersections of embodiment and sensation, culture, and multimodality, and despite developments towards integrating sound analyses in the broad spectrum of the DH and tools attending to sonic collections, scholars such as Lingold *et al.* have been led to echo the sentiment that 'sound remains perhaps the least utilized, least studied mode within DH. Few projects and fewer

tools incite scholars to listen' (Lingold *et al.*, 2018, p. 10).

From a DOH perspective, audio analysis techniques are even more conspicuously and perhaps paradoxically absent. Not diminishing the (oral) historian's duty to engage in source criticism, oral history (OH) often does not seek to validate or gatekeep memory in line with a binary nexus of 'true' or 'false'. This is because——while not foreclosing the ability of OH to recover or restore to the historical record of what was lost, excluded, and devalued—OH places special emphasis on understanding how, why, when, and even whether memories are recalled and filtered through the interview narrative, itself a co-construction of interviewee and interviewer, anchored in the time and space of the interview (e.g. Allen, 1992). OH does this because it is from such runways that analyses of the deeper significance of historical events to individuals and groups, and vice versa, can take flight, potentially opening deeper understandings of how the interchange of culture, memory, interiority, and subjectivity shape perceptions and portrayals of the past (e.g. Passerini, 1979, p. 104; Portelli, 1981, pp. 99–100; Thomson, 2007, p. 53).

Towards such ends, aurality has particular, if underutilized resonance given the window it can open onto the dynamics of meaning-making that can be obfuscated by textual transcripts. As Portelli, for example, has argued, vectors of speech, like pitch and tone, rhythm, and volume shift are beyond the communicative reach of the transcript yet 'carry implicit meaning and social connotations' (2006, p. 34). Changes of rhythm, for example, can occur multiple times in an interview, and may point to the ways that interviewees are framing and narrativizing their recollections: 'Regular grammatical pauses tend to organize what is said around a basically expository and referential pattern, whereas pauses of irregular length and position accentuate the emotional content, and very heavy rhythmic pauses recall the style of epic narratives' (Portelli, 2006, p. 34). So too, the interplay of such vectors, also elided by the transcript, may be significant: Portelli muses that when a narrator slows down (or shifts their speech 'velocity'), they may wish to impart greater emphasis to what they are communicating, or struggle to recall (2006, p. 34). Thus, 'Analysis of changes of velocity must be combined with rhythm analysis' (2006, p. 34).

And if the transcript proves insufficient regarding the capture of such vectors it can prove overdeterminative in other regards: 'Punctuation indicates pauses distributed according to grammatical rules: each mark has a conventional place, meaning, and length. These hardly ever coincide with the rhythms and pauses of the speaking subject' (Portelli, 2006,

p. 35). Summing up the importance that OH places on identifying and theorizing markers of interiority and subjectivity, and the insufficiency of the transcript alone in supporting this, Portelli reflects also on those 'Traits which cannot be contained within segments are the site (not exclusive but very important) of essential narrative functions: they reveal the narrators' emotions, their participation in the story, and the way the story affected them. This often involves attitudes which speakers may not be able (or willing) to express otherwise or elements which are not fully within their control' (2006, p. 35). This emphasis on the necessity of moving beyond the segmentation of the interview, here used by Portelli in the sense of its structural segmentation, will be extended below, as this article calls for a moving beyond the normative, hierarchical segmentation of the OH interview record as transcript versus audio artefact.

Thomson's (2007) article described the ongoing digital revolution as potentially the fourth paradigm transformation of OH. Whilst acknowledging contemporary debates on the practical, ethical, and analytical potential of the digital on OH he noted that the discipline was at that time on the cusp of change and the exact future direction and appearance of digital and digitized OH practice remained uncertain. Despite this uncertainty, Thomson made some prescient predictions about future developments including embracing the possibilities of webcams and remote interviewing to shrink the world and even the development of 'sophisticated digital indexing and cataloguing tools—perhaps assisted in large projects by artificial intelligence—[which] will enable anyone, anywhere to make extraordinary and unexpected creative connections within and across OH collections, using sound and image as well as text' (Thomson, 2007, p. 68). Drawing on Frisch's (2006) concept of a 'post-documentary sensibility' and of digital technology enabling a return to aurality, Thomson (2007, pp. 68–69) identified the potential of searches based on sound and image instead of or more likely in addition to text-based indexing and categorizing to offer new possibilities for searching, analysing, interpreting, and re-presenting oral histories within and across interviews.

In this article, we contend that despite the promise of Thomson's and Frisch's future scanning a recalibrating of aurality in the past decade, speech and sound analysis in OH has remained largely underexplored, opening up rich possibilities for a departure we call Multimodal DOH (MDOH). As this article contends, MDOH can take us beyond more passive notions of 'oral history in the digital age' and the time is ripe to begin thinking about MDOH as a generative subfield of interest in its own right. As observed elsewhere, not only could such a sub-field open new analytical and interpretative

strategies—that are neither founded upon nor predominately focused upon text—for engaging with AV materials, it could contribute to a more thorough and sustained reassessment of the dominance, even tyranny, of the 'written' word in fields like DH and OH (Nyhan and Flinn, 2014).

In this article, we seek to conceptualize MDOH as a 'trading zone' (e.g. Galison, 1996; Kemman, 2021), or a site of interchange and knowledge production that will accrete from and through sustained interdisciplinary encounters: 'Interdisciplinarity is at the very heart of the concept of the trading zone: the transfer and exchange of concepts, methods, tools, techniques, and skills between or across different disciplinary fields or domains' (Fickers and van der Heijden, 2020). Our aim is not, however, to present an ethnographic or observational study of a situated trading zone; instead, we aim to imagine the MDOH trading zone and to follow the questions that its imagining entails: what digital tools are available to and being used by oral historians to access, process, present, search, and analyse interviews? What techniques and tools are being tested and developed, and what may be possible with regards to audio analysis techniques? The DH is equally wedded to the textual record; however, a budding community of scholars is utilizing AV materials, and advancing tools and techniques. While the movement towards 'sound as data' and sound analysis is in its relative infancy (in these fields), it is undoubtedly taking shape. How is the field of DH in particular bridging with DOH? Beyond the clear Venn diagram with digital history and DH, what potential interfaces exist with other fields (e.g. linguistics, music/sound studies, computer science, and AI)? Can we repurpose technology in new ways, and can we scale up such techniques across corpora? With the advancement of audio analysis software, what kinds of humanistic questions might we be able to ask? Ethical issues have also arisen from the transformation to digital, from re-use, re-analysis, and re-interpretation (e.g. Larson, 2013; Crossen-White, 2015; Moravec, 2017) so there must also be a consideration of the ethical challenges that may follow a turn to sound as data, an imperative tied to digital hermeneutics and digital tool criticism.

What follows is therefore primarily concerned with audio as the focal point of analysis within this wider multimodal trading zone. Accordingly, we proceed with a methodological note on the literature search and environmental scan that underpins this article; we seek to summon the MDOH trading zone by identifying its continuities with and disjunctures from more traditional DOH through a literature review and environmental scan; we then seek to plot a course towards MDOH, beginning with the conceptual, methodological, and technical bridging points of existing disciplinary crossovers concerned with computational analysis of spoken word recordings. From there, we move to discuss scholarship that speaks directly to the MDOH research agenda set out here, tracing some of the tools, techniques, and methodological innovations that can help to pave the way for this renewed approach to DOH. From this largely methodological and technical orientation, in the closing sections of this article, we turn to the new research agenda that can emerge from the MDOH trading zone, reflecting also on issues like digital hermeneutics and ethics that ought to be accounted for in the 'sound as data' turn.

1.1 Mechanics of the literature search and environmental scan

An initial survey of the literature, and an environmental scan of websites, and digital tools were carried out in May and June 2020 by students in the University College London DH MA/MSc programme. A literature search was carried out across the major citation and library-search databases to which University College London (UCL) has institutional access, including JSTOR, Google Scholar, Web of Science, and UCL Explore Literature and using n-grams like 'digital AND oral history', collated into a Zotero group folder, and thematically grouped, that is, Social Network Analysis and Text Mining; Data Visualization; and Audio Analysis. These were added to and updated in early 2021 through additional keyword searches for example, 'sound analysis AND oral history', keyword searches of DH journals, and snowballing relevant citations. From this, it was found that research on the digital *analysis* of spoken word collections is being undertaken in the DH and computational sciences primarily, using or developing techniques like automatic speech recognition (ASR), emotion detection, and machine learning, in tandem with automatic transcription and textual methods. Reflecting this, publications and conference proceedings on such digital methods tend to appear in journals across information science, DH, technology, computational linguistics, music studies, and 'big data', rather than established OH journals or websites. Much of this work is still experimental and appears in conferences, workshops, and the work of special interest groups.

The environmental scan of digital tools carried out in May—June 2020 and extended in 2021 focused on three areas—collections management, transcription, and text analysis—and over thirty tools were identified. These were based on recommendations found on the OH websites Digital Omnium, Centre for Oral History and Digital Storytelling, Oral History and Technology, Groundswell-Oral History for Social Change, The Oral Historian's Digital Toolbox, and the digital repositories Zenodo, GitHub, and Figshare.

These were later added to and updated in 2021 to reflect new findings and any decommissioned websites (see Appendix A). The scan, which included social media channels, also returned thirty-one examples of OH project websites, mainly in the English language, for example, The Hong Kong Heritage Project, in order to ascertain the types of technologies used and funding sources in presenting and curating OH collections online.

2 The peculiarities of DOH as currently conceived

What differentiates DOH from traditional OH that uses or benefits from digital technology? One end of the spectrum entails the instrumental use of technology, to do the same things faster or more cost-effectively, like 'democratising' access through the online publication of OH collections or the reduction of manual labour involved in processing interviews. Here the use of ASR technology to expedite transcription is fast encroaching. Next, there is the design and implementation of (semi-)automated and automated tools and methods to allow individual OH interviews and interview collections to be discovered, searched, and interlinked in new ways, usually entailing the maximization of the content of the interview transcript or information about the interview as an object in a collection with metadata. Then there is the emerging and more exploratory 'sound as data' approach or the enfolding of the computational analysis and theorization of audio (and AV) content in both OH and DH circles.

MDOH, we contend, arises in and may extend beyond the sound as data turn, and is a re-orientation of the digital from passive to generative and self-reflexive in the human–machine study of spoken word recordings. It extends DOH as digitally available collections and digital aids to traditional methods and towards DOH *as praxis and unit of analysis*. It is concerned with data-driven analysis, potentially scalable, of OH collections, including understanding (digital) sonification as an analytical category. MDOH is no less concerned with the culturally and socially determined aspects of oral histories from which we might find and interpret meaning; however, analysis and interpretation are to a far greater extent underpinned by computational *processes and reflexivity*, digital tools and methodologies, and a call for engagement with OH artefacts in all their representational modalities: transcript, sound, waveform, metadata. Indeed, what new or interdisciplinary questions might, and might not, arise in such an MDOH framework is of equal interest.

MDOH, as we position it in this article, emphasizes the re-centring of aurality and analytic methods, tools, and processes that attend specifically to the interview as sonic entity, the cultural and subjective aspects of aurality, and the intersubjectivity that may be studied therein. Yet, MDOH's proposed re-centring of aurality and shift in analytic focus exists within a complex framework of established and emerging digital methods for the study of spoken word recordings that, as will be demonstrated, remain heavily indebted to the affordances of the transcript. Thus, in seeking to recentre aurality, to understand it as more than simply additive, we emphasize the potential of a multimodal data-driven analysis and the new understandings of the interconnectedness of OH representational formats it entails.

2.1 DOH: the state of the art

There have been many more books, chapters, and papers that situate 'oral history in the digital age' (or the 'digital revolution') than there is space for in this review, with many being, to the 2022 reader, quaint reviews of the possibilities of the social web and digital dissemination. Broadly, the possibilities for DOH have been tied to a wider shift towards AV sources, a shift that has accelerated in light of the multimedia social web as well as the mass digitization of AV collections previously in analogue formats (Schrum *et al.*, 2012; Warren *et al.*, 2013; Scagliola and de Jong, 2014; Arnold *et al.*, 2021), representing the 'fourth paradigm shift' in OH, that is, the digital turn (Thomson, 2007).

Within the OH community proper, digitality has primarily been lauded for affording greater access and sharing of interview recordings through web 2.0 technologies, improved cataloguing and metadata affordances, indexing, and speech segmentation, information retrieval, and for creating efficiencies in the management and processing of OH interviews overall. With this has come the acknowledgement that OH has been from its beginnings tied to the ability to record interviews using automation technology, and that interview recordings are themselves highly underexploited as data sources, described as 'notoriously underutilized' by Frisch in 2014 (2014, p. 103). Frisch went on to state that the 'content of these collections is rarely organized, much less indexed in any depth, and the actual audio or video is generally not searchable or browseable in any useful way', let alone provision for sound-led organization or analysis (Frisch, 2014, p. 103).

Have these calls heralded advances in making the aurality of the interview computationally tractable? Certainly, major advances have been made in indexing and making audio and audio-video collections searchable, coming under the umbrella of 'oral history digital indexing' (Lambert, 2019a) through work by Frisch, Lambert, and others. In 2014, Frisch touched upon

many of the technological developments and future possibilities for DOH that have since born fruit, such as searchability and interactivity of digitized audio/visual collections, time-coding or segmentation, annotation (tending towards standardization) across large-scale collections and relational databases, AI, and emotion mapping/tagging. And though these utterances around audio led Freund to claim an 'attack' on the transcript, warranting a re-assertion of its value (Freund, 2017, p. 33), audio and AV have not unseated the textual record of the interview. While re-centring voice is the essence of such projects as Audio-Video Barn (Warren *et al.*, 2013) most agree there is still little interest in audio and video despite being paramount to all of OH (e.g. Walker and Halvey, 2017) and comparatively fewer projects implement technology for advanced audio analysis. Audio (as does video) seems to be perceived as additive to the transcript, allowing context and body language not captured in the transcript to be recovered. Systematic or computational analysis of the interview primarily via the medium of sound is rare.

For the most part, work that situates itself in 'digital oral history' is still heavily reliant on textual transcriptions (Boyd, 2013). Clement *et al.* observed that most content management tools available for audio and visual materials, whether open-source or proprietary, had hitherto been 'designed to leverage human-generated transcripts and metadata and do not provide any means for analysing the audio itself' (2014, p. 2)—summarizing the very lacuna of interest to this review. In addition, there is a general lack of awareness of what technology is available to researchers, archivists, and librarians working with or managing OH collections (Walker and Halvey, 2017). Stakeholders still placed the value of AV material squarely in engagement and dissemination, and in bolstering the authenticity of the user experience (Walker and Halvey, 2017). Even more telling, perhaps, was the differential between text versus audio as reflected in research versus dissemination: 'The transcript was recommended for research purposes, whilst audio and visual engagement were considered to be useful for engaging younger demographics and wider members of the public' (Walker and Halvey, 2017, pp. 1294–5).

The digitization and structural and semantic enrichment of the transcript have undoubtedly brought OH closer to the DH. However, there remains a vast gap between an OH community that is interested in digital possibilities for their work under the umbrella of 'digital oral history' and sound-oriented digital humanists, computer scientists, sociolinguists, and computational linguists. The former is concerned with preservation, curation, access (e.g. Matusiak *et al.*, 2017), dissemination, 'democratising', information

retrieval, transcription, ethics, OH pedagogy, 'representing history in visual forms' (Lambert, 2019b), the usability of digital resources (Boyd and Larson, 2014), and facilitating communication and collaboration (High, 2010; Thomson, 2016). The latter, with sound as data, the possibilities for bringing advanced technologies and DH methods to bear upon spoken word and multimodal recordings and using technologies to ask new or expanded humanistic questions. The following section outlines some of the predominant tools and methods employed by oral historians but also scholars using spoken word corpora as data.

Preservation, collections management, and, above all, access, are foundational concerns of DOH (Lambert and Frisch, 2013; Matusiak *et al.*, 2017) and a multitude of platforms for working with AV materials are recommended by oral historians, particularly Omeka but also ArchivSpace, Collective Access, Islandora, and Dédalo (see Appendix A). Their affordances include metadata creation, indexation, file combination, file validation, curation of collections and exhibitions, geo-referencing/mapping and visualization, and some such as Dédalo support multilingual interpretation. A range of desktop applications and web-based platforms, most are out-of-the-box while others require more advanced technical knowledge. Other tools available to and recommended by oral historians specifically for processing and preserving AV files include Fixity, BFW MetaEdit, and Exactly. All underpinned by a free or freemium model, they are part of a suite of tools created by Audio Visual Preservation Solutions, which 'was founded based on an identified need to help organizations tackle the challenges and build capabilities to access and preserve their audiovisual collections' (Audio Visual Preservation Solutions, 2017). BWF MetaEdit (About this Software, no date), for example, is open-source and facilitates embedding, validating, and exporting of metadata in Broadcast WAVE Format (BWF), CSV, and XML, using 'checksums' to mitigate audio file corruption. Developed with support from the Library of Congress and the US Federal Agencies Digitization Guidelines Initiative (FADGI), it was designed for audio archivists, documentarians, and collections managers (Audio Visual Preservation Solutions, 2017a).

Closely tied to current, text-centric modes of accessing and interacting with OH collections are the advancement of transcription and information retrieval techniques. Information retrieval for improving access to OH collections remains a foundational approach, such as through indexing (Boyd, 2013; Gref *et al.*, 2018), segmentation (Gustman *et al.*, 2002), and the use of ASR and machine learning to make recordings searchable (Salesky *et al.*, 2016). The use of video has also been strongly tied to promoting greater

engagement with OH archives, and as an enhancement to access and 'digital storytelling', including in education (Christel *et al.*, 2010; High, 2010; Kaufman, 2013; Gould and Gradowski, 2014). Most of the literature surveyed on the use of digital methods for analysing OH or spoken word collections focuses on various text-based analyses such as social network analysis and text-mining (McKether *et al.*, 2009; Verd and Lozares, 2014, McKether and Friese, 2016; O'Reagan and Fleming, 2018), semantic networking (Pattuelli and Miller, 2015), topic modelling and tagging, and data-visualization (Ohno *et al.*, 2010; Paju *et al.*, 2011; Xiao *et al.*, 2013).

Sociolinguistic and natural language processing studies also make use of OH corpora for dialectology and the localization of ASR for lesser-used languages through text–speech alignment and normalization (Braber and Davies, 2016; Scherrer *et al.*, 2019). ASR does not deal with people speaking over one another, nor background noise, nor accents and dialects, and always requires correction, whether manually or digitally assisted. With training, accuracy rates may be improved, but this may require a significant investment of time and technical expertise. Indeed, for the creators of the OH Metadata Synchronizer (OHMS), indexing was originally a way to overcome costly and labour-intensive transcription on a large scale, and the technical limitations of ASR (Boyd, 2013).

OHMS remains one of the leading resources that have been created by and for oral historians. Free, open source, and web based, this software aims to enhance interaction between users and OH collections, in audio and video form, and was created by the Louie B. Nunn Centre for OH at the University of Kentucky Libraries (Oral History Metadata Synchronizer, no date). Originally designed 'as an archival context to serve an archival access imperative' (Boyd, 2013, p. 96), it improves online access through indexing, transcription synchronization (i.e., matching transcribed text to time-codes in the recording in order to 'connect the user from a location in the text to a moment in the recording' (Boyd, 2013, p. 97)), and meta-data management. It provides word-level search capability and a time-correlated transcript or indexed interview connecting the search term to the corresponding utterance and moment in the recorded interview online. Though aimed primarily at improving access, interaction, and information retrieval, it additionally deploys alternative entry points for information seeking like emotion tagging and emotion detection (Warren *et al.*, 2013; Truong *et al.*, 2014; Turner, 2017). Other transcription and text analysis tools, both free and commercial, cited by oral historians include WebASR, FromTo (see below), Express Scribe, NVivo, ATLAS.ti, GATE, Lexalytics, and Apache Open Natural Language Processing (NLP) to name a few.

2.2 Bridging points: from DOH towards MDOH

What then are the most promising avenues for wrangling the 'non-linear' affordances of the recording (Frisch, 2014, p. 113), and for engaging with the OH interview as sonic, digital artefact? What disciplinary interfaces hold potential for, or are mutually beneficial with, MDOH? Indeed, the consequences of developing MDOH are not limited to OH and heritage collections, just as spoken word recordings are not solely circumscribed by OH.

Linguistics in the first instance seems a natural bridging point. This is suggested due to what we hold to be the ongoing, if implicit, re-contextualization, and re-evaluation of the dominance of the transcript that is in process in several linguistics projects and extending also into the data-driven and algorithmic approaches being used in linguistics, for example, ASR. ASR is pervasive across the recent literature given its potential for reducing transcription labour, but it is also now playing a role in audio analysis, annotation, and segmentation, bridging the transcript to the audio, acting as both a conceptual and computational connector of these formerly segmented surrogates. In essence, then, if a suggestion to 'put aside' the transcript and the dominance of the written work in OH that it entails is not only a theoretical over-correction, but a technological impossibility given the role of markup and other forms of annotation in making computational objects machine readable, then our call is for the repositioning and interweaving of the transcript as one layer among many.

In the field of sociolinguistics, the use of an OH corpus has been demonstrated in regional dialect research (Braber and Davies, 2016). Braber and Davies, demonstrating the utility of legacy collections and reworking underused collections, used existing OH archives (the East Midlands Oral History Archive) and created new recordings through community outreach for intergenerational comparison. Though employing traditional methods, the study draws attention to the need to scale up sociolinguistic studies to avoid 'cherry-picking' methods that bias the findings and extrapolations. Underlining this, it contemplates the 'secondary' use (Braber and Davies, 2016, p. 100) and analysis of OH archives and associated data beyond their original purpose and the ethical and methodological issues this poses, a point we will return to below. The focus on dialect, the authors argue, also constitutes a return to the origins of OH in folklore and dialectology and is therefore mutually beneficial to both oral historians and sociolinguists (Braber and Davies, 2016).

Others still argue that there is much to gain from closer contact between linguists and oral historians. Both stand to gain from the growing banks of recordings and text corpora preserved, restored, and digitized by national and international entities as well as a long-mooted return to 'the "oral" in oral history' (Roller, 2015), an additional crossover being the study of endangered and first nations languages and dialects (Kasten *et al.*, 2017). Further still, corpus linguistics has been advocated for the analysis of oral histories for multilingual collections (Pagenstecher, 2018), while Roller has applied the technique to study aspects of the Welsh-English dialect over time in addition to historical insights regarding the decline of the coal mining industry using transcribed interviews from national collections (Roller, 2015). Reflecting this intersection of linguistics, OH, and DH, one tool designed for computational linguistic approaches to spoken word recordings is ELAN, an open-source 'linguistic annotator' (Transcription, no date) for AV recordings. ELAN allows users to create layered annotations through tiers and tier hierarchies, supporting controlled vocabularies and the linking of multiple video files in one annotation document. ELAN is maintained by The Language Archive (TLA) whose remit is linguistics and which holds multi-lingual spoken word collections including 'endangered and understudied' languages, as well as recorded Sign language collections (The Language Archive (ELAN), no date). Thus, the post-textual paradigm at play in these projects entails transcription, and layers of textually imposed intervention in the form of markup expressed in natural language codes that structurally connect the machine readable-transcript with its AV compatriot. In contrast with a more traditional OH or passive DOH paradigm, the transcription arguably does not have precedence here, but it is one data layer of many, one entrance point of many, alongside but not in hierarchical opposition to the AV surrogate of the captured event as it once occurred in time and space.

ASR, besides expediating transcription, is also recognized as a tool that fundamentally promotes access—a way of making collections and the data they contain more searchable and discoverable. Commercial (e.g. Dragon Speech Recognition) but also research-led ASR tools (e.g. WebASR, University of Sheffield) are increasingly becoming available for aiding transcription of OH interviews at the same time as researchers working with spoken word recordings are turning to machine learning and the advancement of automated transcription. This has been indispensable in supporting audio analysis in humanistic scholarship. Staying with linguistics, some have aimed their sights at ASR for multilingual OH archives (Byrne *et al.*, 2004), improving ASR for lesser known languages or dialects through combined acoustic modelling techniques to improve indexing, transcription, and curation, and instrumentalizing OH archives and their metadata as test datasets (Gref *et al.*, 2018, 2019; Gorisch *et al.*, 2020).

3 'Tooling up' for MDOH

It is, however, now possible to move even beyond such bridging points and disciplinary interfaces, to methods and projects that are within touching distance of the MDOH paradigm explored here.

Significant ongoing research is reworking OH recordings and re-centring aurality with techniques like machine learning to enhance discovery (Clement *et al.*, 2014), emotion detection, and analysis. Crucially, much of this research is moving away from an overemphasis on the textual transcript of the interview, by looking at pitch, vocal effort, and pause duration in the speech signal with a view to navigating collections by emotion markers (Truong *et al.*, 2014). Four initiatives that have discerned the advanced possibilities for DOH, directly and indirectly, are AudioVisual Material in DH (AVinDH); CLARIN Media Suite; the Sussex Humanities Lab (SHL); and HiPSTAS.

AvinDH is a special interest group (SIG) in the exchange of knowledge and methods for working with narrative-based audio, video, and visual material across multiple disciplines including OH and linguistics (Home, no date). The group is motivated by the questions of how to analyse spoken word recordings 'as sound rather than as (just) content', and what tools can facilitate this, at scale, and with multiple data types (Background, no date). The group recognizes that:

> Audiovisual sources have a potentially huge value for the Digital Humanities... A single document can provide information regarding language, emotions, speech acts, narrative plots and references to people, places and events.

> ...no widespread accepted digital research methodology for the exploration and analysis of audiovisual content in a data-driven manner exists as of yet. In fact, data modelling based on the study of patterns and trends that can be discerned in the various semantic layers of audiovisual (sub)collections has only just begun to gain attention (Background, no date).

Having reviewed the AvinDH workshops at the annual DH conference since 2014, Stefania Scagliola, a founding member of this SIG, has stated that papers showcasing technological innovation have dominated but that some 'focused less on tool development and more

on the analysis of content, like the study of changes in voices in news coverage...' (Scagliola, 2021, p. 12).

A recent special issue of *Digital Humanities Quarterly* resulted from the work of AvinDH (Arnold *et al.*, 2021) and focuses predominantly on film and music data sources as well as environmental sounds. This covers image sonification (Kramer, 2021), Afrofuturism and music remix (Thompson and Carerra, 2021), and sonification of data for augmented reality (Bonnett *et al.*, 2021). Two contributions are of particular interest to MDOH. First, Have and Enevoldsen's use of 'convolutional neural networks' (CNNs) on spectrograms to ascertain, on a large scale, the proportion of music to speech in radio programming over the last quarter-century using archived Danish radio programmes. Though the focus is to aggregate different aspects of 'speech-music discrimination' (ratio of speech to music) the authors further suggest that this model could be trained to detect gender or regional accents (Have and Enevoldsen, 2021).

Another contribution seeks to 'demonstrate the potential of the computational methods to hear gentrification, using a combination of ethnography, passive acoustic recording, and computational sound analysis' (Martin, 2021, p. 4). Martin's study is concerned with the policing of blackness and 'black sound' through silencing in US streetscapes and neighbourhoods and responding to the #dontmuteDC movement (Martin, 2021, p. 6), much of the sound data being derived from recordings of a traffic intersection in Washington DC. However, it uses a mixture of interviews and soundscape recordings that were cleaned and processed using clustering and classification algorithms (in Kaleidoscope and Fast Fourier Transform, respectively) as well as manual tagging of sounds, training the dataset, and uploading to an SQL database for analysis. Echoing the work of Clement, Webb *et al.*, and Boyd, Martin similarly recognizes the lateness of DH in treating sound as data and endeavouring to 'leave behind the text as an anchoring force, and to create work with large sonic datasets that respects the call and parameters of black digital humanities' (Martin, 2021, p. 10). Also striking in Martin's methodology is a rigorous grounding in humanistic interpretation, using various sonic representations and 'computationally informed close listening' (Martin, 2021, p. 23) facilitated by spectrograms (using Google Chromes Music Lab Spectrogram tool). A microcosm of the traffic intersection, Martin visualizes the sound of sirens over the period recorded as a way of thinking about the relationship between policing, blackness, urban space, and gentrification. Less about spoken word and more about the contingencies of multiple sound sources, this study is instructive in its methods and humanistic anchoring.

Common Language Resources and Technology Infrastructure (CLARIN) is an EU-funded initiative with consortiums across European universities and research institutions. For example, FiNER, the Finnish NER tool used by Kannisto and Kauppinen (above) is part of the Fin-CLARIN consortium at the University of Helsinki. CLARIN-ERIC is compiling a database of existing OH collections across Europe (Existing OH Collections in Europe, no date). There are annual conferences and workshops on research using, for example, the CLARIN Media Suite, which is part of the Dutch infrastructure for DH (Skadin and Eskevich, 2018) and works closely with the Dutch 'Sound and Vision' archive. While the CLARIAH Media Suite has provided some user-friendly ASR, transcription, and segmentation resources aimed at oral historians, its full capabilities remain highly specialized, a problem traversed by many studies reviewed here.

All of the CLARIN workshops are reported at Oral History & Technology (OH&T), a website that summarizes in an accessible way, the principal digital techniques available and used by oral historians and provides guides and manuals on available techniques and resources. Additionally, the 'OH Portal' facilitates speech recognition, transcription correction, word alignment, and phonetic detail, thus it 'presents an audiovisual display of the audio signal' (Oral History, no date). The platform is multilingual and free for academic use (requiring institutional sign-in). Among others, it uses 'FromTo', software designed to convert the results of ASR into a 'Karaoke-style' html-page. The resulting HTML-file (with embedded java-script) shows the text and, optionally, the sound file can be played while highlighting the word spoken. It also provides ASR corrector software (FromTo, no date).

The SHL has attempted to address the potentialities of spoken word recordings through the implementation of music information retrieval (MIR) methods on a small sample of use cases (Webb *et al.*, 2017). This study redeployed out-of-the-box toolkits at a DH2016 conference workshop to experiment with alternative analyses of spoken word data namely visualizing audio data, inspecting and extracting audio segments and features, mapping textual information to sound to identify and visualize sections of similarity or changes in speech, and identifying speaker-voice characteristics (e.g. gender). Crucially, the authors suggest that the latter may be scalable and portable, able to enrich metadata or create it bottom-up for collections that are not yet annotated or easily explored. SHL has also created Old Bailey Voices using the Old Bailey Online dataset (and others), a 'working "macroscope"' for scaled exploration of both 'big data' and a single datum and combining multiple data sources (tim.benskitchen.com). Another similarly structured SHL creation for

the BBC is an OH 'workbench', also a development in multimedia and macroscope technology using OH recordings as well as transcripts and metadata. Together these demonstrate the possibilities for multi-modal access, interaction, and analysis, and for scaling up across corpora and different types of data. Elsewhere, MATLAB has also been used for audio signal analysis of spoken word recordings including MIR among a suite of techniques including feature extraction, representation, and filtering, audio classification, segmentation, and audio sequence alignment (Giannakopoulos and Pikrakis, 2014). Though this route requires technical knowledge, it demonstrates that there are established tools and digital environments that can accommodate multiple audio-analysis modalities that can potentially be repurposed, emulated, or remixed for MDOH.

The High-Performance Sound Technologies for Access and Scholarship project (HiPSTAS), a web-based platform (HiPSTAS, no date) offers a series of open-source tools using Adaptive Recognition with Layered Optimization (ARLO) software for creating and sharing AV annotation. It also has capabilities in audio searching, spectral visualization, matching, classification, clustering, audio-tagging, and audio labelling for machine learning (Clement *et al.*, 2014; Clement and McLaughlin, 2017). A web-based, machine-learning application, ARLO was repurposed from its origins in acoustic studies in animal behaviour and ecology. It uses spectrograms, that is, visual representations of audio files (Webb *et al.*, 2017), 'to model sonic features for machine learning processes including clustering (unsupervised learning) as well as classification (prediction or supervised learning)' (Clement, 2016, p. 15). A major use case of the HiPSTAS project has been spoken word poetry collections. It has been used in searching poetry collections by sound in order to assess 'vocal gestures' in spoken poetry performances, allowing scholars to 'consider new research questions about the performances including a focus on sonic "para-content" such as pitch and laughter' (Clement, 2016, pp. 16–17) as well as other interactive features of performances, like applause (Clement and McLaughlin, 2016). As the authors suggest, identifying and quantifying various audio features can open up several humanistic questions and be transposed onto different phenomena, in this case, performance and 'the cultural context of poetry' (Clement and McLaughlin, 2016, p. 10). However, the project casts its eye over 'audiovisual heritage' much more broadly to include OH collections, and with recent Mellon Grant funding is in collaboration with several other major US archives and research centres such as the Louie B. Nunn. According to Clement, the goal of the project workflow is to 'broaden the use of audiovisual

materials in research, teaching, and the public by scaling up the use of international standards and free audiovisual annotation frameworks and making AV, ultimately, easier to find and use' (UT News, 2020).

In designing HiPSTAS, Clement *et al.* incorporated an assessment of user requirements into the development of ARLO for analysing large and diverse sound collections, experimenting with laughter and applause identification. Through this consultation and experimentation process, the authors reported that oral historians entered with intentions for text-based content extraction and left with an expanded desire to address how sound analysis could be used to address questions of embodiment, gender, intercultural relations, and how visualization of interviews might transcend the transcript—themes that were similarly explored by Webb *et al.* (2017). Other users refined their research questions to a much higher degree of specificity after experimenting with ARLO, reflecting also the limitations of the software itself and what could be asked of it. Experimenting with MIR techniques, Webb *et al.* (2017) also reported an expanded sense of possibilities for DH uses of OH collections among the more computationally literate participants but much less so in novices. Nevertheless, they also observed that after workshopping the tools, concepts, and documentation, many participants 'felt inspired to rethink their current forms of analysis and to investigate how they might incorporate new forms of computational analysis' (Webb *et al.*, 2017, p. 180) as well as coming around to the idea of working collaboratively with data scientists.

These examples demonstrate the potential of awareness building to transform thinking and practice around DOH and the problem of the digital skills gap which raises questions about how these advanced techniques can be taken up by the OH community. This point has similarly been raised by Walker and Halvey in assessing UK-based OH technologies and their use by different stakeholders (2017) begging the questions: who might these advanced computational tools or techniques for audio/AV analysis serve? Are they for the OH community to adopt? Are they ultimately for 'secondary' use in analyses of collections and data by DH scholars or truly aiming at trans-disciplinary development? While open-source software, such as for ASR, may still require expert skills to set up and navigate (Ordelman and van Hessen, 2018, p. 165), tools like Transcription Portal (Transcription Portal, no date), as well as commercial providers, are bridging this gap.

4 New Questions

What kinds of new questions might oral historians, digital humanists of various disciplinary shades, and others be able to ask and answer with the development

of digital, audio analytical tools, and methods, and with appropriate awareness of available and developing technologies? Salah and Pessanha have taken emotion detection a step further and explored how the computational analysis of non-verbal cues and biological signals in OH interviews—breaks, gaps, gestures, facial expressions, silences, and breathing patterns—allow us to study feelings, mood, and culture at scale. This experimental work, using machine learning and AI tools on trauma-related speech corpora, suggests pathways for enriching analyses of the subjectivities of memory, storytelling, and historical trauma. They also explore the protection of informants' identities in politicized environments, while using computational tools to safely make available emotion or trauma-related OH data (Pessanha and Salah, 2021; Multimodal Digital Oral History, 2022a).

The History of Emotions is one area of historical research where the potential of an MDOH approach may be articulated. The History of Emotions has become, over the past 20 years, a thriving and interdisciplinary sub-field of the discipline of History. Interested in the emotions of groups and individuals alike, the field understands emotions as potentially playing an important role in historians' aim of 'enlightening us on the historical meaning of what people really experienced according to their own worldview, explaining motivation and seeking specific results in their own words' (Nagy, 2019, p. 189).

What are emotions? How can they be studied? And how are they captured by and in historical and born-digital sources? The field continues to struggle with ontological, epistemological, and methodological questions around the identification, definition, and analysis of emotion. A key problem is in identifying residual traces of emotions in historical sources that represent aspects of an otherwise elusive phenomenon—that may be communicated through presence, absence, movement, and action; through tone, velocity, modularity, and tonality of speech and even silence—as an (often linguistic) vestige (Nagy, 2019, pp. 197–200).

The rebalancing of the pre-eminence of the written word, and the re-centring of the oral, aural, and sonic affordances called for by MDOH has the potential to open OH collections to historians of emotion as sites of historical inquiry, allowing not only the communication of emotions in digital sources to be studied in new ways but directing attention to the source criticism of digital sources including their affordances for representing emotions. MDOH may furthermore offer Historians of Emotion a test set through which to research digital approaches: the detection, analysis, and historicization of emotion, in individual interviews but also at scale. As such, MDOH has the potential to make contributions to what Nagy summarized as two

dominant approaches to the History of Emotions: the 'topic' approach, or how a category of emotion is manifested in a given period and the 'anaytical category approach', where emotions may offer a way into the analysis of wider social and cultural phenomenona. If 'working on past emotions is particularly prone to evoke emotions; the point is brought to the forefront to an even greater extent for those doing oral history or other forms of history in the present' (Nagy 2019, p. 199), then MDOH can enhance such reflexivity, potentially offering a new meeting ground for interdisciplinary investigations into the coding and decoding of emotions, how they may be defined and studied and how they are shaped by digital tools and their affordances.

Sound Studies (Sterne, 2012), like linguistics, offers a rich corpus of literature attending to the socio-cultural entanglements of sound over time, what and how we hear, and 'the sonic technologies that mediate and construct our experiences' (Lingold *et al.*, 2018, pp. 4–5). This suggests some crossover with fields like media archaeology and may be of particular interest to DOH that handles older, analogue recordings that have subsequently been digitized, or sonic 'artefacts' (Lingold *et al.*, 2018, p. 4) of previous technologies. Published in 2018, *Digital Sound Studies* explores the intersections and future pathways of sound studies and DH, and the ways in which audio and visual analyses are being used and challenging academic paradigms of knowledge transmission. Placing 'sounds in their cultural, historical, and social contexts' (Lingold *et al.*, 2018, p. 5), chapters engage music and soundscapes, both contemporary and historical, as well as the creation of multimedia archives and collaborative research. The editors also emphasize humanistic interpretation and that close reading need not be sacrificed in the use of digital tools, echoing the hybridity of DH (van Zundert, 2015).

Most relevant is Clement's elaboration of the epistemological challenges of sound classification using the aforementioned ARLO software for HiPSTAS. Machine learning for recognizing and tagging qualities of speech requires initial human input of 'seed' data, itself a highly subjective process (Clement, 2018). Clement reiterates that underlying the difficulties of training a system to classify speech elements is the humanistic reality that 'how we perceive and make meaning with prosodic and paralinguistic features is a subjective activity' (p. 163). Prosody is concerned with intonation, stress, and rhythm and 'can signify elements of a speaker's identity including affect and emotional engagement, age, cognitive process and development, ethnicity, gender, and region' (p. 163); paralinguistics with the much more difficult task of compartmentalizing such interpreted sounds as laughs,

sighs, hesitations or crying, in other words, the semi-verbal facts of speech (Scagliola and de Jong, 2014). Given what she describes as the 'emergent nature of sound hermeneutics in digital space' (2018, p. 169) as opposed to the more established realm of text, Clement reasons that adaptation is required to balance philosophical, technical, and practical concerns in our desire to represent and conceptualize sound data in its true complexity: shifting from familiar, text-oriented tools to those 'change-of-paradigm' (Clement, 2018, p. 165) tools that can accommodate dynamic sonic metadata; 'fixed' meanings to multiplicity, ambiguity, and 'emergent' (Clement, 2018, p. 168) interpretations across different disciplinary boundaries; and discrete to relational and contextual representations of speech/sound features.

Elsewhere, Clement's work on 'dissonant records' is premised on close listening to resistance in historical audio, such as in the collections of the American Folklife Centre at The Library of Congress, to reveal socio-technical distortions in oral histories. She reflects on the challenges that audio distortions, like interpreted sounds, pose for computational analysis, and centres audio as resonant artefacts and 'opportunities for understanding the process of history as one that is deeply entangled in politics and media' (Multimodal Digital Oral History, 2022b).

Envisioning systems attuned to the layered, relational, contextual, and emergent meanings of sound, will entail some reduction, and Clement (2018) insists upon an expansive and hermeneutical approach to technological sound analysis. Indeed, in this very insistence, the potential theoretical productiveness of MDOH is suggested, including its potential to bridge back to and enrich the DH itself, given the emphasis being put on understanding the hermeneutical entanglements and consequences of the field (van Zundert, 2015).

4.1 Digital hermeneutics and the ethics of MDOH

If decontextualization is an inevitable part of the digital (Pagenstecher, 2018), digital tools also carry the potential for misrepresentations antithetical to OH values. Being at once a methodological and ethical concern, here our discussion strays into digital hermeneutics, which is concerned with the ways in which technologies (tools, algorithms, information systems, and visualizations) both interpret and impact the interpretation process, and the novel interpretations they might herald (Romele et al., 2020). Or as it has been defined with a view to digital history and humanities, 'the critical and self-reflexive use of digital tools and technologies for the development of new research questions, the testing of analytical assumptions and the production of

sophisticated scientific interpretations' (Fickers and van der Heijden, 2020).

From this directly stems digital source criticism, in the sense of an inquiry into the decisions and processes (technical, social, institutional, financial, cultural, epistemological, etc.) that have influenced how an OH collection has been retro-digitized or born digital. In MDOH, questions of selection, context, representation, ownership, and access are all crucial to a digital hermeneutically inflected source criticism as they directly shape the kinds of questions that can usefully and ethically be asked of both retro-digitized and born-DOH collections. Yet also in play is 'digital tool criticism', which is the evaluation of the digital tools we use to conduct research, their limitations and compatibility with the research question(s), and the ways in which they shape the entire research process (Koolen et al., 2019, pp. 381–382). In such a framework, a hermeneutics of MDOH cannot be 'post-processing' or 'post-algorithmic' (van Zundert, 2015, pp. 355, 342) but must be part of the life-cycle of a research project. This begs the question as to whether or how we can accommodate the practices and philosophy of 'close-listening' in large-scale data-driven audio analyses of OH collections, that is, a 'hermeneutics of in-betweenness' (Fickers and van der Heijden, 2020).

Given the relative novelty of the terrain and the emergent nature of technologies in its service, digital source, and tool criticism should be a key facet of MDOH, as should accounting for the ethics of digital access and data-driven analyses. Indeed, we maintain that the 'methodological issue of interpreting with digital machines' (Romele et al., 2020, p. 74) is itself an inherently ethical concern, given that 'sound has a politics; it can be gendered and racialized, used both to liberate people from and reinscribe determinative social categories. Sound has ethical implications...' (Lingold et al., 2018, p. 5).

In this regard, as we have elsewhere suggested in the context of cultural heritage more broadly, bringing both computational analysis and a feminist DH lens to bear upon the spoken record of the past would be a powerful alignment. Already given the pervasiveness of cultural heritage data in DH, MDOH has much to gain from engagement with feminist DH. Feminist DH is predicated on critiquing the reproduction of oppressive structures in DH research paradigms and digital environments more broadly, and seeks to rebuild them ethically and from an intersectional feminist standpoint (Smyth et al., 2020). Traditionally, OH has itself been heavily inflected by feminism and feminist ethics of care and collaboration (Sheftel, 2018). Likewise, emerging from feminist DH are calls for the care and repair of feminist projects (Losh and Wernimont, 2018), not least DOH archives and tools, and for

ongoing evaluation of consent, ownership, re-interpretation, and re-use. Not simply a question of mitigating the shortcomings—and misrepresentations—of the transcript through greater access to recordings of women's and marginalized voices (Larson, 2018), Sheftel reminds us that a 'feminist ethic of digital oral history' needs to account for both the voices that want to be heard and those that do not in the abundant Internet (2018, p. 281).

The affordances of the social web have in this sense raised questions about the renewal of consent for interviews primarily because new sharing and analytical possibilities are not necessarily desirable on the part of interviewees. This creates challenges around reuse beyond the original purpose and agreements of the OH collection and associated data, control over dissemination, open-data research funding models, representation and (potentially) manipulation of recordings, and the limitations of what is legal and what is ethical in interpreting copyright transfer from an interviewee(s) to a repository (Larson, 2013; Gluck, 2014; Braber and Davies, 2016; Sheftel and Zembrzycki, 2017).

Such a feminist ethics of care thus links up with the wider questions that scaling up, making recordings available, and multimodal analysis may demand. ASR services, for example, are commonly web-based meaning that, as the authors of OH&T caution, any data uploaded will be stored or passed via an external server, risking exposure of any sensitive data in an uploaded interview. Web-based services tend, however, to be much easier to use than desktop software, which also requires more processing power (Oral History, no date).

Furthermore, ASR technology is underwritten by machine learning and AI and for which there may be additional disparities based on regional dialect or accent, and whose efficacy is further contingent on environmental factors, and the degree to which interviewers/-ees speak over one another (Gorisch *et al.*, 2020). AI tools indeed have a particular tendency to reproduce racial and gender biases (Gebru , 2020) a fact that is well documented across DH and digital cultural heritage literature (Noble, 2018; Risam, 2018). Freund has further brought home how DOH has become caught up in technologies of surveillance and surveillance capitalism and re-iterates the need to historicize modalities (Multimodal Digital Oral History, 2022c). In this, we can include automation and algorithmic tools in the service of OH. Additionally, they are not, as Jones and Edendberg posit, one-size-fits-all resources and are thus poorly understood in their specificity and complexity, reducing participants' ability to make informed choices around consent (2020). If we are to practice 'responsible AI' (Dignum, 2020, p. 215) for MDOH, layered over the longer-running concerns

regarding consent and re-use outlined above, due diligence must be paid to the human input that underpins such tools (Dignum, 2020) and research methodologies as well as risk and bias assessment, accountability, and transparency.

Complementing the academic literature, MDOH may do well to draw on, localize and perhaps even problematize or extend, the various guidelines emerging from or tailored to national and trans-national data contexts. Recommended guidelines include the UK's 'Understanding artificial intelligence ethics and safety' (GOV.UK, no date b), the EU-issued *Ethics-guidelines for trustworthy AI* (Publications Office of the European Union, no date) and those recommended by the OECD AI Policy observatory (OECD, no date), the Ada Lovelace Institute (Ada Lovelace Institute, no date), and the Alan Turing Institute public policy initiative (The Alan Turing Institute, no date) to mention but a few.

UK government-issued guidance, alongside the Data Ethics Framework (GOV.UK, no date a) suggests that establishing ethical building blocks for an AI project is crucial. In particular, it should be verified that an AI project is 'ethically permissible', 'fair and non-discriminatory', 'worthy of public trust', and 'justifiable'. To support action as well as reflection, guidance also calls for the drawing up and use of the following instruments: 'Framework of ethical values', 'Set of actionable principles', and 'Process based governance framework' (see GOV.UK, no date b). These frameworks could be pursued by MDOH through sustained and reflexive attention to the human-, machine- and data foundations and imbrications not only of the digital tools and infrastructures that the field would seek to work with but also with regards to the retro-digitized and born-digital interview corpora it would aggregate or generate and further utilize.

Space will not allow further discussion of these instruments here, yet they clearly represent a crucial next step in the development of MDOH and are open questions for future research.

5 Conclusion

The 'the long view' of OH, from its origins in the folklorist tradition, emphasizes its long relationship with technology (Martin, 2021). In this article, OH is once more at a crossroads. While text still dominates, emergent sound analysis along with the digital's remediation of heterogeneous formats into a common binary format suggests a need to define OH anew for the data-driven context. Thanks to a growing number of tools, studies, as well as major scholarly and institutional initiatives to preserve, digitize, and manipulate 'sound heritage' (broadly conceived) for research, it is possible

to speak of an MDOH as we have begun to conceptualize it in this article. HiPSTAS, CLARIAH Media Suite, AVinDH, and the SHL represent some of the most developed thinking and building in this pursuit. Nevertheless, quality metadata, transcription, or some other form of manual and textual intervention remain prerequisites in most case studies that delve into speech recognition and audio analysis SIG techniques. Thus, we suggest that the aim cannot be to truly detach from the transcription or natural-language-based code intervention in our aspirations for this. Rather, it is for MDOH to reconceptualize the textual as one layer of many, a layer that is interconnected, co-constitutive, relational, and accorded parity of esteem with all other layers, whether human or machine derived.

Acknowledgements

Funding

References

About this Software. (no date). BFW MetaEdit. http://bwfmetaedit.sourceforge.net/ (accessed 20 August 2021).

Ada Lovelace Institute. (no date). Ada Lovelace Institute. https://www.adalovelaceinstitute.org/ (accessed 23 March 2022).

Allen, B. (1992). Story in Oral History: Clues to Historical Consciousness. *The Journal of American History*, **79**(2): 606–611.

Arnold, T., Scagliola, S., Tilton, L., and Gorp, J. (2021). Introduction to special issue on audio visual data in DH. *Digital Humanities Quarterly*, **15**(1).

Audio Visual Preservation Solutions. (2017). About US, AVP. https://www.weareavp.com/about-us/ (accessed 20 August 2021).

Audio Visual Preservation Solutions. (2017a). BFW MetaEdit, AVP. https://www.weareavp.com/bwf-metaedit/#bwf-download (accessed 20 August 2021).

Background. (no date). AVinDH SIG: special interest group audiovisual material in digital humanities. avindhsig.wordpress.com/background/(accessed 10 August 2021).

Bonnett, J., Bolton, J., Ralph, W., Legault, A., MacAfee, E., Winter, M., Jacques, C., and Anderson M. (2021). Annotating our environs with the sound and sight of numbers: the dataScapes project. *Digital Humanities Quarterly*, **15**(1).

Boyd, D. (2013). OHMS: enhancing access to oral history for free. *Oral History Review*, **40**(1): 95–106.

Boyd, D. A. and Larson, M. A. (eds) (2014). *Oral History and Digital Humanities: Voice, Access, and Engagement*. New York, NY: Palgrave Macmilan.

Braber, N. and Davies, D. (2016). Using and creating oral history in dialect research. *Oral History*, **44**(1): 98–107.

British Library. (2019). Unlocking our sound heritage. https://www.bl.uk/projects/unlocking-our-sound-heritage# (accessed 29 July 2021).

Byrne, W., Doermann, D., Franz, M., Gustman, S., Hajic, J., Oard, D., Picheny, M., Psutka, J., Ramabhadran, B., Soergel, D., Ward, T., and Zhu, W. (2004). Automatic recognition of spontaneous speech for access to multilingual oral history archives. *IEEE Transactions on Speech and Audio Processing*, **12**(4): 420–35.

Christel, M. G., Stevens, S., Maher, B., and Richardson, J. (2010). Enhanced exploration of oral history archives through processed video and synchronized text transcripts. *Proceedings of the 18th ACM International Conference on Multimedia*, Firenze Italy, October 25–29, 2010, p. 1333.

Clement, T. E., Tcheng, D., Auvil, L., and Borries, T. (2014). High performance sound technologies for access and scholarship (HiPSTAS) in the digital humanities. *Proceedings of the American Society for Information Science and Technology*, **51**(1): 1–10.

Clement, T. E. (2016). Towards a rationale of audio-text. *Digital Humanities Quarterly*, **10**(2): 1–14.

Clement, T. E. (2018). Word spoken articulating the voice for high performance sound technologies for access and scholarship (HiPSTAS). In Lingold, M. C., Mueller, D., and Trettien, W. (eds), *Digital Sound Studies*. Durham: Duke University Press, pp. 155–77.

Clement, T. and McLaughlin, S. (2016). Measured applause: toward a cultural analysis of audio collections. *Journal of Cultural Analytics*, **86**(1): 1–17.

Clement, T. and McLaughlin, S. (2017). Introducing the HiPSTAS audio toolkit workflow: audio labeling, HiPSTAS. http://hipstas.org/2017/08/31/introducing-the-hipstas-audio-toolkit-workflow-audio-labeling/ (accessed 10 August 2021).

Crossen-White, H. L. (2015). Using digital archives in historical research: what are the ethical concerns for a 'forgotten' individual? *Research Ethics*, **11**: 108–19.

Dignum, V. (2020). Responsibility and artificial intelligence. In Dubber, M. D., Pasquale, F., and Das, S. (eds), *The Oxford Handbook of Ethics of AI*. New York, NY: Oxford University Press, pp. 213–31.

Existing OH Collections in Europe. (no date). Oral history & technology. http://oralhistory.eu/collections/clarin-eric (accessed 10 August 2021).

Fickers, A. and van der Heijden, T. (2020). Inside the trading zone: thinkering in a digital history lab. *Digital Humanities Quarterly*, **14**(3).

Freund, A. (2017). From.wav to.txt: why we still need transcripts in the digital age. *Oral History*, **45**(1): 33–42.

Frisch, M. (2006). Towards a post-documentary sensibility: theoretical and political implications of new information technologies in oral history. In Perks, R. and Thomson, A. (eds), *The Oral History Reader*, 2nd edn. London: Routledge, pp. 102–14.

Frisch, M. (2014). Oral history and the digital revolution: towards a post-documentary sensibility. In Boyd, D. A. and Larson, M. (eds), *Oral History and Digital Humanities:*

Voice, Access, and Engagement. New York, NY: Palgrave Macmillan, pp. 102–14.

FromTo. (no date). Oral history & technology. https://oralhistory. eu/oh-tools/fromto#download (accessed 10 August 2021).

Galison, P. (1996). Computer Simulations and the Trading Zone. In P. Galison and DJ Stump (eds) *The Disunity of Science: Boundaries, Contexts, and Power*. Stanford: Stanford University Press, 118–157.

Gebru, T. (2020). Race and gender. In Dubber Markus D., Pasquale, F., and Das, S. (eds), *Oxford Handbook of Ethics of AI*. Oxford: Oxford University Press, pp. 253–69.

Giannakopoulos, T. and Pikrakis, A. (2014). *Introduction to Audio Analysis: A MATLAB Approach*. 1st edn. Oxford: Academic Press.

Gluck, S. B. (2014). Reflecting on the quantum leap: promises and perils of oral history on the web. *The Oral History Review*, **41**(2): 244–56.

Gorisch, J., Gref, M., and Schmidt, T. (2020). Using automatic speech recognition in spoken corpus curation. Proceedings of the 12th Language Resources and Evaluation Conference, Marseille, France, 13-16 May 2020. *ELRA - European Language Resources Association*. http://www.lrec-conf.org/proceedings/lrec2020/index.html#6423 (accessed 3 February 2021).

Gould, J. G. and Gradowski, G. (2014). Using online video oral histories to engage students in authentic research. *The Oral History Review*, **41**(2): 341–50.

GOV.UK. (no date a). Data ethics framework. https://www.gov. uk/government/publications/data-ethics-framework (accessed 23 March 2022).

GOV.UK. (no date b. Understanding artificial intelligence ethics and safety. https://www.gov.uk/guidance/understanding-artificial-intelligence-ethics-and-safety (accessed 23 March 2022).

Gref, M., Schmidt C., Behnke, S., and Kolher, J. (2019). Two-staged acoustic modeling adaption for robust speech recognition by the example of German oral history interviews. *2019 IEEE International Conference on Multimedia and Expo (ICME)*, Shanghai, China, 8–12 July 2019, pp. 796–801. doi: 10.1109/ICME.2019.00142.

Gref, M., Köhler, J., and Leh, A. (2018). Improved transcription and indexing of oral history interviews for digital humanities research. *LREC 2018: Eleventh International Conference on Language Resources and Evaluation*, 7-12 May 2018, Miyazaki, Japan. http://www.lrec-conf.org/proceedings/lrec2018/index.html (accessed 3 February 2021).

Gustman, S., Soergel, D., Oard, D., *et al.* (2002). Supporting access to large digital oral history archives. *The Second ACM/IEEE-CS Joint Conference*. New York, NY: ACM Press, p. 18.

Have, I. and Enevoldsen, K. (2021). From close listening to distant listening: developing tools for speech-music discrimination of Danish music radio. *Digital Humanities Quarterly*, **15**(1).

High, S. (2010). Telling stories: a reflection on oral history and new media. *Oral History*, **38**(1): 101–12.

HiPSTAS. High Performance Technologies for Access and Scholarship (no date). http://hipstas.org/ (accessed 22 July 2021).

Home. (no date). AVinDH SIG: special interest group audio visual material in digital humanities. avindhsig.wordpress.com/(accessed 10 August 2021).

Howes, D. (ed) (2021c). Geography and anthropology. First issued in paperback. In Howes, D. (ed), *Senses and Sensation: Critical and Primary Sources*, Vol. **1**. London: Routledge.

Howes, D. (ed) (2021d). History and sociology. First issued in paperback. In Howes, D. (ed), *Senses and Sensation: Critical and Primary Sources*, Vol. **2**. London: Routledge.

Howes, D. (2014). Introduction to sensory museology. *The Senses and Society*, **9**: 259–67.

Jones, M. L. and Edenberg, E. (2020). Troubleshooting AI and consent. In Dubber, M. D., Pasquale, F., and Das, S. (eds), *The Oxford Handbook of Ethics of AI*. New York, NY: Oxford University Press, pp. 357–74.

Kasten, E., Roller, K., and Wilbur, J. (eds) (2017). *Oral History Meets Linguistics*. Eyüpsultan, Istanbul: Verlag Der Kulturstiftung Sibirien.

Kaufman, P. B. (2013). Oral history in the video age. *Oral History Review*, **40**(1): 1–7.

Kemman, Max. (2021). Trading Zones of Digital History. De Gruyter Oldenbourg 2021.

Koolen, M., Van Gorp, J., and Van Ossenbruggen, J. (2019). Toward a model for digital tool criticism: reflection as integrative practice. *Digital Scholarship in the Humanities*, **34**(2): 368–85.

Kramer, M. J. (2021). What does A photograph sound like? Digital image sonification as synesthetic audiovisual digital humanities. *Digital Humanities Quarterly*, **15**(1).

Lambert, D. (2019a). International oral history workshop, centre for contemporary and digital history. https://www.c2dh.uni.lu/thinkering/international-oral-history-workshop (accessed 18 August 2021).

Lambert, D. (2019b). Oral history as a multimedia and multidimensional presentation challenge. *26th Berlin EVA*. Berlin: Luxembourg University. https://orbilu.uni.lu/handle/10993/42535 (accessed 8 February 2021).

Lambert, D. and Frisch, M. (2013). Digital curation through information cartography: a commentary on oral history in the digital age from a content management point of view. *Oral History Review*, **40**(1): 135–53.

Larson, M. (2013). Steering clear of the rocks: a look at the current state of oral history ethics in the digital age. *Oral History Review*, **40**(1): 36–49.

Larson, M. A. (2018). The medium is political and the message is personal: feminist oral histories online. In Srigley, K., Zembrzycki, S., and Iacovetta, F. (eds), *Beyond Women's Words: Feminisms and the Practices of Oral History in the Twenty-First Century*. London: Routledge, pp. 298–303.

Lingold, M. C., Mueller, D., and Trettien, W. (2018). Introduction. In Lingold, M. C., Mueller, D., and Trettien, W. (eds), *Digital Sound Studies*. Durham, NC: Duke University Press, pp. 1–25.

Losh, E. and Wernimont, J. (eds) (2018). *Bodies of Information: Intersectional Feminism and Digital Humanities*. Minneapolis, MN: University of Minnesota Press.

Martin, A. (2021). Hearing change in the chocolate city: computational methods for listening to gentrification digital humanities and black sound. *Digital Humanities Quarterly*, **15**(1).

Matusiak, K. K. *et al.* (2017). Finding access and digital preservation solutions for a digitized oral history project: a case study. *Digital Library Perspectives*, **33**(2): 88–99.

McKether, W. L. and Friese, S. (2016). Qualitative Social Network Analysis with ATLAS.ti: Increasing Power in a Black Community. ATLAS.ti User Conference 2015 – qualitative data analysis and beyond / Editor: Susanne Friese and Thomas Ringmayr; Universitatsverlag der TU Berlin. https://doi.org/10.14279/DEPOSITONCE-5152.

Mckether, W. L., Gluesing, J. C., and Riopelle, K. (2009). From interviews to social network analysis: an approach for revealing social networks embedded in narrative data. *Field Methods*, **21**(2): 154–80.

MDOH. (2022a). The forward-view with almila akdag salah, francisca pessanha, and myriam fellous-sigrist. Multimodal digital oral history. https://multimodaldigitaloralhistory.omeka.net/exhibits/show/blog/06-july-2022 (accessed 14 September 2022).

MDOH. (2022b). The forward-view with Tanya clement. Multimodal digital oral history. https://multimodaldigitaloralhistory.omeka.net/exhibits/show/blog/22-june-2022 (accessed 14 September 2022).

MDOH. (2022c). The forward-view with douglas lambert and Alexander Freund multimodal digital oral history. https://multimodaldigitaloralhistory.omeka.net/exhibits/show/blog/-1--8-june-2022 (16 September 2022).

Moravec, M. (2017). Feminist research practices and digital archives. *Australian Feminist Studies*, Vol. **32**. Abingdon: Routledge, pp. 186–201.

Nagy, P. (2019). History of emotions. In Tamm, M. and Burke, P. (eds), *Debating new Approaches to History*. London: Bloomsbury Academic, pp. 208–17.

Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism, How Search Engines Reinforce Racism*. New York, NY: New York University Press.

Nyhan, J. and Flinn, A. (2014). Oral history, audio-visual materials and digital humanities: a new 'grand challenge'? In *Sound and (Moving) Image in Focus Pre-Conference Workshop. Digital Humanities Conference*. Switzerland.

O'Reagan, D. and Fleming, L. (2018). The FinFET breakthrough and networks of innovation in the semiconductor industry, 1980–2005: applying digital tools to the history of technology. *Technology and Culture*, **59**(2): 251–88.

OECD. (no date). Artificial intelligence, OECD.org. https://www.oecd.org/digital/artificial-intelligence/ (accessed 23 March 2022).

Ohno, S., Saito, S., and Inaba, M. (2010). A platform for mining and visualizing regional collective culture. In Ishida, T. (ed), *Culture and Computing*. Berlin, Heidelberg, Germany: Springer Berlin Heidelberg, pp. 188–99.

Oral History. (no date). Oral history & technology. oralhistory.eu (accessed 5 May 2021).

Oral History Metadata Synchronizer. (no date). Oral History Metadata Synchronizer. https://www.oralhistoryonline.org/ (accessed 5 May 2021).

Ordelman, R. and van Hessen, A. (2018). Speech recognition and scholarly research: usability and sustainability. In Skadin, I. and Eskevich, M. (eds), *CLARIN Annual Conference 2018*. Paris, France: PISA, pp. 163–8.

Pagenstecher, C. (2018). Testimonies in digital environments: comparing and (de-)contextualising interviews with holocaust survivor Anita Lasker-Wallfisch. *Oral History*, **46**(2): 109–18.

Paju, P., Malmi, E., and Honkela, T. (2011). Text mining and qualitative analysis of an IT history interview collection. In Impagliazzo, J., Lundin, P., and Wangler, B. (eds), *History of Nordic Computing 3*. Berlin, Heidelberg, Germany: Springer Berlin Heidelberg, pp. 433–43.

Paju, P., Oiva, M., and Fridlund, M. (2020). Digital and distant histories: emergent approaches within the new digital history. In Fridlund, M., Oiva, M., and Paju, P. (eds), *Digital Histories: Emergent Approaches within the New Digital History*. Helsinki, Finland: Helsinki University Press, pp. 3–18.

Passerini, L. (1979). Work, ideology and consensus under Italian fascism. *History Workshop*, **8**: 82–108.

Pattuelli, M. C. and Miller, M. (2015). Semantic network edges: a human-machine approach to represent typed relations in social networks. *Journal of Knowledge Management*, **19**(1): 71–81.

Pessanha, F. and Salah, A. A. (2021). A computational look at oral history archives. *Journal on Computing and Cultural Heritage*, **15**(1): 1–16.

Portelli, A. (1981). The peculiarities of oral history. *History Workshop*, **12**: 96–107.

Portelli, A. (2006). What makes oral history different. In Perks and Thomson (ed.) *The Oral History Reader*. Second Edition. USA and Canada: Routledge, pp 32–42.

Publications Office of the European Union. (no date). Ethics guidelines for trustworthy AI. https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1 (Accessed 23 March 2022).

Risam, R. (2018). What passes for human? Undermining the universal subject in digital humanities. In Losh, E. and Wernimont, J. (eds), *Bodies of Information: Intersectional Feminism and Digital Humanities*. Minneapolis, MN: University of Minnesota Press, pp. 39–56.

Roller, K. (2015). Towards the "oral" in oral history: using historical narratives in linguistics. *Oral History*, **43**(1): 73–84.

Romein, C. A., Kemman, M., Birkholz, J., Baker, J., de Gruijter, M., Meroño-Peñuela, A., Ries, T., Ros, R., and Scagliola, S. (2020). State of the field: digital history. *History*, **105**(365): 291–312.

Romele, A., Severo, M., and Furia, P. (2020). Digital hermeneutics: from interpreting with machines to interpretational machines. *AI and Society*, **35**(1): 73–86.

Salesky, E., Ray, J., and Shen, W. (2016). Operational assessment of keyword search on oral history. In *10th International Conference on Language Resources and Evaluation (LREC)*. LREC, Portorož, Slovenia, pp. 317–321.

Scagliola, S. (2021). Founding the special interest group audio-visual in digital humanities: an interview with Franciska de Jong, Martijn Kleppe, and Max Kemman. *Digital Humanities Quarterly*, **15**(1).

Scagliola, S. and de Jong, F. (2014). Clio's talkative daughter goes digital: the interplay between technology and oral accounts as historical data. *The Making of the Humanities, Volume III: The Modern Humanities*. Amsterdam, Netherlands: Amsterdam University Press, pp. 511–26.

Scherrer, Y., Samardžić, T., and Glaser, E. (2019). Digitising Swiss German: how to process and study a polycentric spoken language. *Language Resources and Evaluation*, **53**(4): 735–69.

Schrum, K., Brennan, S., Halabuk, J., Leon, S., and Scheinfeldt, T. (2012). Oral history in the digital age. *The Oxford Handbook of Oral History*, Oxford: Oxford University Press, pp. 1–19.

Sheftel, A. (2018). Introduction to section 5: listening to and learning from stories in the digital world. In Srigley, K., Zembrzycki, S., and Iacovetta, F. (eds), *Beyond Women's Words: Feminisms and the Practices of Oral History in the Twenty-First Century*. London: Routledge, pp. 279–82.

Sheftel, A. and Zembrzycki, S. (2017). Slowing down to listen in the digital age: how new technology is changing oral history practice. *The Oral History Review*, **44**(1): 94–112.

Skadin, I. and Eskevich, M. (eds) (2018). *CLARIN Annual Conference 2018*. 8-10 October 2018, Pisa, Italy.

Smyth, H., Nyhan, J., and Flinn, A. (2020). Opening the "black box" of digital cultural heritage processes: feminist digital humanities and critical heritage studies. In Schuster, K. and Dunn, S. (eds), *Routledge International Handbook of Research Methods in Digital Humanities*. London: Routledge, pp. 295–308.

Sterne, J. (ed.) (2012). *The Sound Studies Reader*. London: Routledge.

The Alan Turing Institute. (no date). Public policy. https://www.turing.ac.uk/research/research-programmes/public-policy (accessed 23 March 2022).

The Language Archive (ELAN). (no date). The language archive (ELAN). https://archive.mpi.nl/tla/ (accessed 10 August 2021).

Thompson, T. L. and Carerra, D. (2021). Afrofuturist intellectual mixtapes: a classroom case study. *Digital Humanities Quarterly*, **15**(1).

Thomson, A. (2007). Four paradigm transformations in oral history. *Oral History Review*, **34**(1): 49–70.

Thomson, A. (2016). Digital aural history: an Australian case study. *Oral History Review*, **43**(2): 292–314.

Transcription. (no date). Oral history & technology. https://oralhistory.eu/oh-tools/transcription (accessed 10 August 2021).

Transcription Portal. (no date). Speech data & technology. https://speechandtech.eu/oh-portal (accessed 11 April 2022).

Truong, K. P., Westerhof, G., Lamers, S., and de Jong, F. (2014). Towards modeling expressed emotions in oral history interviews: using verbal and nonverbal signals to track personal narratives. *Literary and Linguistic Computing*, **29**(4): 621–636.

Turner, K. (2017). Creating history: a case study in making oral histories more accessible in the digital age. *Digital Library Perspectives*, **33**(1): 48–62.

UT News. (2020). Mellon grant helps UT austin scholars preserve and promote audiovisual heritage, UT News. https://news.utexas.edu/2020/10/07/mellon-grant-helps-ut-austin-scholars-preserve-and-promote-audiovisual-heritage/ (accessed 10 August 2021).

Verd, J. M. and Lozares, C. (2014). Reconstructing social networks through text analysis: from text networks to narrative actor networks. In Dominguez, S. and Hollstein, B. (eds), *Mixed Methods Social Networks Research*. Cambridge: Cambridge University Press, pp. 269–304.

Walker, I. and Halvey, M. (2017). On designing an oral history search system. *Journal of Documentation*, **73**(6): 1281–98.

Warren, R. E., Maniscalco, M., Schroeder, E., Oliver, J., Huitt, S., Lambert, D., and Frisch, M. (2013). Restoring the human voice to oral history: the audio-video barn website. *The Oral History Review*, **40**(1): 107–25.

Webb, S., Kiefer, C., Jackson, B., Baker, J., and Eldridge, A. (2017). Mining oral history collections using music information retrieval methods. *Music Reference Services Quarterly*, **20**(3–4): 168–83.

Xiao, L., Luo, Y., and High, S. (2013). CKM: a shared visual analytical tool for large-scale analysis of audio-video interviews. *2013 IEEE International Conference on Big Data*. Piscataway, NJ: IEEE, pp. 104–12.

van Zundert, J. J. (2015). Screwmeneutics and hermenumericals: the computationality of hermeneutics. *A New Companion to Digital Humanities*. Hoboken, NJ: Wiley-Blackwell, pp. 331–47.

Appendix A.

Tools (software, code sources) surveyed

Name	URL	Use	Access
Apache OpenNLP	https://opennlp.apache.org	Text analysis	Free, Open Source
ArchivesSpace	https://archivesspace.org	Collections management	Free, Open Source
ATLAS.ti	https://atlasti.com/	Text analysis	Paid, Freemium model, Proprietary
BWF MetaEdit	http://bwfmetaedit.source forge.net/	Collections management	Free, Open Source
Cloud	https://cloud.google.com/	Text analysis, Transcription	Paid, Proprietary
CollectiveAccess	https://www.collectiveaccess.org/	Collections management	Free, Open Source
CONTENTdm	https://www.oclc.org/en/con tentdm.html	Collections management	Paid, Proprietary
Dédalo	https://dedalo.dev/	Collections management	Free, Open Source
DiscoverText	https://discovertext.com/	Text analysis	Paid, Proprietary
Dragon Speech Recognition	https://www.nuance.com/ dragon.html	Transcription	Paid, Proprietary
Drupal	https://www.drupal.org	Collections Management	Free, Open Source
ELAN	https://archive.mpi.nl/tla/elan	Transcription	Free, Open Source
Exactly	https://www.weareavp.com/ products/exactly	Collections Management	Free, Open Source
Express Scribe	https://www.nch.com.au/ scribe/index.html	Transcription	Free, Proprietary
Fixity	https://www.weareavp.com/ products/fixity-pro	Collections Management	Paid, Proprietary
FromTo	https://oralhistory.eu/software/ fromto	Transcription	Free, Open Source
GATE	https://gate.ac.uk/	Text Analysis	Free, Open Source
Historypin	https://www.historypin.org/en/	Collections Management	Free, Proprietary
Islandora	https://islandora.ca/	Collections Management	Free, Open Source
Joomla	https://www.joomla.org/	Collections Management	Free, Open Source
Lexalytics	https://www.lexalytics.com/	Text Analysis	Paid, Proprietary
MAXQDA	https://www.maxqda.com/	Text Analysis	Paid, Proprietary
MediaInfo	https://mediaarea.net/en/ MediaInfo	Collections Management	Free, Open Source
NVivo	https://www.qsrinternational. com/nvivo-qualitative-data-analysis-software/home/	Text Analysis	Paid, Proprietary
NYPL	https://github.com/NYPL/oral-history	Collections Management	Free, Open Source
OHMS (Oral History Metadata Synchronizer)	https://www.oralhistoryonline.org/	Collections Management	Free, Open Source
Omeka	https://omeka.org	Collections Management	Free, Open Source
Oral History Timeline Templates	https://github.com/titanium bones/Oral-History-Timeline-Templates	Collections Management	Free, Open Source
Otter.ai	https://otter.ai/	Transcription	Paid, Fremium model. Proprietary
Transana	https://www.transana.com/	Transcription	Paid, Freemium model, Proprietary
WebASR	https://www.webasr.org/	Transcription	Paid, Freemium model, Proprietary
WordPress	https://wordpress.org/ download/	Collections Management	Free, Open Source