

Standards, Vokabulare und KI-Technologien für die qualitative Datenanalyse von Oral-History-Interviews

Katrin Moeller

1. Oral History und digitale Transformation

Mittlerweile kann die Geschichtswissenschaft auf eine Vielzahl von Sammlungen und Dokumenten der Oral History zurückgreifen, die zu unterschiedlichsten Themen, Zeiten und Aspekten angelegt wurden. Mit dem Portal „Oral History.Digital“ (oh.d) werden solche Bestände nun auch für eine größere Forschungscommunity gemeinsam sichtbar und vor allem direkter erreichbar. Dies verstärkt gleichzeitig den Wunsch, große Bestände noch besser integriert zu erschließen und mit komplexen Fragestellungen übergreifend auszuwerten. Reizvoll ist etwa die Idee, nach Vorbild des Deutschen Referenzkorpus (DeReKo) oder des Deutschen Textarchivs (DTA) für die Sprach- und Literaturwissenschaften einen gemeinsam und tief erschlossenen Forschungskorpus für Oral History zu schaffen. Auf dem Weg dorthin sind jedoch viele praktische Herausforderungen der Datenkuratation zu überwinden. Mit der stark wachsenden und sich intensiv vernetzenden Forschungscommunity, dem Portal oh.d und vielen daran angeschlossenen Geschäftsstellen und Agenturen sowie dem Konsortium NFDI4Memory ist die Community auf dem Weg dahin.

Dieser Beitrag möchte dazu digitale Transformationsprozesse analysieren, die Oral History als eine besondere Quellengattung der Geschichtswissenschaft erfassen und digitale Arbeitsprozesse und Herausforderungen systematisieren. Dies erfolgt, indem die bisherigen Phasen der Digitalisierung systematisch betrachtet und aktuelle Herausforderungen beschrieben werden. Besonderes Augenmerk soll dabei auf die Verwendung von Vokabularen und Normdaten, heute auch oft übergreifend als Linked Open Data bezeichnet, gelegt werden, die zur automatisierten Transkription, Erschließung, digitalen Repräsentation, Datenanreicherung und Analyse im Prozess der Digitalisierung eine völlige neue Rolle bekommen haben. Es soll gezeigt werden, wie solche Arbeitsmittel die digitale Verarbeitung unterstützen.

2. Ausgangssituation

Oral History war bisher als Methode der Geschichtswissenschaft wie auch im Kontext der bürgerschaftlich orientierter Geschichtswerkstätten ein Arbeitsfeld ausgewiesener Expertinnen und Experten, die vor allem hermeneutische, qualitative Methoden zum Einsatz brachten und deren Hemmschwelle zum Umstieg auf digitale Methoden (mit meist algorithmenbasierten und damit quantitativ-statistischen Analyseverfahren) besonders hoch einzuschätzen ist (Franken 2020). Dies ist auch Produkt einer besonders

in den Anfangsjahren der Oral History geäußerten vehementen Kritik aufgrund der gegenseitigen Abhängigkeit von Forschungsergebnissen, Fragestellungen und Interviewtechniken (Leh 2023: 59 ff.). Gegen solche Kritik konnte sich die Oral History behaupten, weil sie sich schnell eine dezidierte hermeneutische Methodik mit Regularien für Arten von Fragestellungen, Formen der Interviewführung und Leitlinien zur Transkription von Texten sowie Methoden erarbeitete, die jeweils von den Forschungsinteressen und projektleitenden wissenschaftlichen Fragestellungen abhängig waren (Wierling 2003; Fuß/Karbach 2019).

Mit der Digitalisierung bahnt sich jetzt ein zweiter großer methodischer wie organisatorischer Umbruch innerhalb der Oral History an, der hier als Ausgangssituation in drei großen Prozessen kurz skizziert werden soll. Dabei sind diese Phasen eher von ihrem Innovationscharakter her zu lesen, nicht so sehr als eingrenzende historische Etappen oder Zeiträume. Jede Innovation durchläuft meist eine längere Phase der Modifikation und Weiterentwicklung, die nicht zwangsläufig nur von einer Technologie oder Verfahrensweise bestimmt wird.

(1) Diese Veränderungen in der Forschungsdisziplin begannen mit den stark erleichterten Aufnahmetechniken durch die massenhafte Verfügbarkeit von zunächst einfacheren und später multimedialen Aufnahmegeäten sowie insbesondere von Handys, die zu einer neuen Popularität von Oral History führten. Mittlerweile entstehen größere Interviewsammlungen nicht mehr nur im Rahmen von Forschungsprojekten mit dezidierten Fragestellungen oder zielgerichteten bürgerbewegten Communities mit spezialisierten Technikausrüstungen, sondern zunehmend innerhalb von Citizen Science und bürgernahen Projekten durch „Jedermann“ oder „Jederfrau“ und eben auch mit einer Vielzahl von Formaten und Methodiken. Gerade Oral History stellt seit ihren Anfängen eine intensive Schnittstelle zwischen Forschung und Citizen Science dar. Ebenso erfolgt die Sammlung von Interviews verstärkt auch angesiedelt an Museen, Gedenkstätten und Gedächtniseinrichtungen, die Oral History als eine zentrale Aufgabe der Erinnerungskultur und des kulturellen Gedächtnisortes für sich entdeckt haben. Damit wird die durch konkrete Fragestellungen oder von Forschungsprojekten getriebene Oral History durch Interviewsammlungen ergänzt, die eher einem breiten Ansatz der Erinnerungskultur folgen und zum Teil eher lebensgeschichtliche Narrative verfolgen. Dies verändert letztlich auch die Inhalte und Analyseformen für solche Dokumente. Gleichzeitig treten durch die Langzeitarchivierung und digitale Sammlung von Beständen Studienergebnisse der älteren Forschung anderer Fachdisziplinen, beispielsweise der Psychologie oder Soziologie sichtbar hervor, die nach und nach für die historische Forschung interessant werden und im oh.d-Portal oder den Repositorien der verschiedenen Fachdisziplinen ihren Aufbewahrungsort finden.

Mit der Vielfalt der technischen Geräte wuchs und wächst zugleich die Mannigfaltigkeit von Aufnahmequalitäten, Dateiformaten und Codecs, was nicht immer mit der Langzeitarchivierung und DFG-Praxisregeln konformgeht (Pagenstecher 2017: 86 f.). Immer noch ist zudem die Rettung von Aufzeichnungen auf älteren Medienträgern (Tonbänder, Kassetten etc.) ein wichtiges Thema. Zugleich werden methodische Konzepte der Oral History neu herausgefordert, weil Bürgerwissenschaftler*innen häufig andere Herangehensweisen und Methoden verfolgen, aber auch neue Wissensfelder besetzen.

(2) Eine weitere Etappe der Digitalisierung formierte sich mit der Etablierung von Forschungsinfrastrukturen und der zunehmend automatisierten Transkription durch