

Artificial Intelligence and Oral History: The Good, the Bad and the Ugly

Jill Cassidy

Report on Dr Doug Boyd's Bernard Bailyn Lecture¹

Jill Cassidy initiated the formation of what is now Oral History Tasmania in 1991 while working for Launceston's Queen Victoria Museum & Art Gallery and, apart from a two-year period, has been its president ever since. She received the Hazel de Berg award in 2015.

As Artificial Intelligence (AI) becomes freely available, it's proving to be increasingly accurate for the transcription of Oral History interviews. But as Boyd points out in this lecture, there are downsides which also need to be considered.

As the archivist at the Louie B. Nunn Centre for Oral History at the University of Kentucky Libraries, Boyd has focused particularly on increasing access to the collection. Troubled by the fact that the audio of an interview was disconnected from the text, he used his technological knowledge to envision, design and implement the Oral History Metadata Synchronizer (OHMS) which synchronises text with audio and video online. It is open source and free and is now used widely around the world. In 2008 the Nunn Centre had 6,000 interviews but only around 200 were accessed in a year. Now using OHMS the rate of access has increased dramatically; in 2023 the collection had 18,000 interviews and were accessed 238,000 times.²

OHMS requires text to search so needs a transcript. For decades oral historians have longed for accurate speech recognition software to transcribe quickly and cheaply, but there have always been problems as machines don't like accents or background noise or people talking over each other. Boyd occasionally tested software but it was

1 The Bernard Bailyn Lecture in North American History for Latrobe University, 21 November 2023.

2 For a discussion of this see Judy Hughes, 'Oral History, Digital Technologies and Innovation: From Interview to Archive', *Oral History Australia Journal* 41 (2019): 80–81. Available at https://oralhistoryaustralia.org.au/wp-content/uploads/2019_journal_80-81_Hughes.pdf. Accessed 28 August 2024.

never perfect. However, it's possible now to have a reasonably good transcript, good enough for searching.

Boyd uses a free system called Whisper (from the people who released ChatGPT) which was better than a human transcriber in several instances. It's actually a library so it has to be installed on a server. However, there is now the app MacWhisper for those with a Mac. Boyd points out that AI can never be as good as a human but Whisper's accuracy is very exciting as the Nunn Centre can now generate text for every interview in the collection, and not just a small proportion with the rest simply indexed. It's important to note that Whisper provides just a block of text which still needs to be formatted and the speakers' names added which Boyd estimates takes 6½ hours per hour of interview. But this opens up a whole new era.

Boyd now also has the ability to go into a transcript's data and extract key words such as people, places, dates and events which is invaluable for searching. Name Entity Recognition (NER) is used, also known as Natural Language Processing. Boyd showed a page from a transcript that had been plugged into an NER engine called spaCy; the various items asked for are highlighted in a different colour for each type: for example, blue for dates, green for people and purple for institutions. Even if the transcript is not perfect, it's possible to extract very good metadata.

There is a potential bias in name recognition. By mid-2025 he hopes to have NER incorporated into OHMS to help eliminate this problem. NER is not perfect. Boyd used the example from a series of interviews about Jim Beam bourbon. Jim Beam was a person but it's also a brand of whiskey and a distillery. He's building in an editor which teaches the program which of these is meant.

So far so good. But as innovation is accelerated so is human risk. Boyd is increasingly being asked for the online interviews to be taken down, and it is their very accessibility which is the problem. In 2019 he was already questioning how we can explain informed consent and how to protect people's privacy. Interviewees have signed consent forms but are now discovering just how accessible their interviews are. For example, in preparation for a job interview, they check what comes up in a search engine and find the interview comes up second and they don't want some of

the deeply personal revelations made public, such as the mention of drugs, sexually transmitted diseases or criminal allegations.

As Boyd points out, most Oral History interviews are about details, about someone's life. He doesn't think anyone is getting informed consent right as it's impossible to explain how the interview will be used when we ourselves don't know what the true ramifications are. Earlier interviewees did not understand just how widely their interview could spread and the limited audience in past decades made the risk very small. The worry is that people today will be more reluctant to do life interviews. Archives might consider them too dangerous and make them available online only after death, although Boyd did not elaborate on how this could be determined. In the meantime, he is considering a click-through to filter out bots. Ten years ago he would have considered this a barrier to access.

Archives need to have informed accessioning to know what is coming into the archive. After producing the AI-generated transcript he intends to build in a sensitivity analysis for interviews, using a list of some 70 terms that might be of concern, such as 'cocaine', 'secret', 'don't record this' and 'abuse'. It doesn't mean they are going to be refused, just that the archive knows what might be problematic in their collection.

Boyd searched his database for 'maiden name' (75 hits), 'elementary [primary] school' (408 hits) and 'best friend' (150 hits). These could be enough to reset a bank password. As he explains, privacy issues preceded AI but AI is intensifying the situation. He questions how we are going to feel if our interviews are being used by ChatGPT.

Generative AI is the ugly side of AI, leading to post-truth and fake history. It's possible to generate something that previously didn't exist with just a few minutes of voice recording using Speech Synthesis. There are many available apps that can be instructed to provide anything required in the voice of whoever audio is available for. Boyd describes it as Photoshop for the voice, and believes it is inevitable that we're going to move into fake history based on fake primary sources. He sees this as

a problem for society as a whole, not just archives and historians, if we can't really believe that everything we see is real.

On the positive side, there is a coalition working already on a Content Authenticity Initiative. The idea is to give audio and video a cryptographic signature to certify its authenticity which will stay with the recording all its life. Boyd also predicts there will be a lot of research done to train AI to detect fakes, and that historians and Oral History archives will have a new role as authenticators, able to prove the authenticity of what they are saying. And the fact that interviews are online does at least point to one way to establish authenticity.

It's worth watching the lecture just to see some of the transcript of an AI-generated 'interview' that Boyd initiated, with Alistair Thomson supposedly the interviewer;³ absolutely fake from start to finish and with ludicrous Australianisms but with a reasonably believable account of changes in fatherhood in the lifetime of the 'interviewee'.

The lecture finished with a nod to the 'good' use of AI in helping facilitate transcriptions in different languages. Whisper can do speech recognition in a great number of languages, including rare ones such as Afrikaans, Albanian and Haitian, so AI might break the English hegemony in transcripts.

3 The lecture can be accessed on YouTube at <https://www.youtube.com/watch?v=DOg0iCefZJw>.