Improving Eye Witness Testimony (EWT) in Air Accident Investigation Through the use of AI Generative Pre-Trained Transformers (GPTs)

ISASI Rudolph Kapustin Memorial Scholarship 2024

Ben Wright

Cranfield University

School of Aerospace, Transport and Manufacturing

MSc in Safety and Human Factors in Aviation

Supervisor: Dr Wen-Chin Li

Improving Eye Witness Testimony (EWT) in Air Accident Investigation Through the use of AI Generative Pre-Trained Transformers (GPTs)

Introduction

The primary focus of air accident investigation is to improve safety outcomes for the wider aviation industry. This is done through an investigative process that prioritises scientific rigour to determine causality through the accurate analysis of evidence. EWT is one stream of evidence that suffers from inherent inaccuracy, leading to a reduction in the quality of the investigative output. In recent years, Artificial Intelligence (AI) has offered intuitive solutions to problems in almost every industry. The following paper applies the AI GPT functionality to propose AI chatbots as an investigative tool within the air accident investigation process to improve the accuracy of EWT recall and the subsequent evidence validation process.

The Challenge of EWT Validation in Investigation

Put simply, EWT - the process of recalling information from a prior event - is unreliable. This is demonstrated by Loftus & Palmer's (1974) seminal study which applied novel scientific methodology to EWT research to illustrate how the phrasing of questions asked by investigators could lead to inaccurate recollections of events witnessed by participants compared to highly-validated events.

Our perception of events can be altered and made more inaccurate by external influences; this is known as the 'misinformation effect'. Factors that can affect the recall of an event include social influences following the event (Gabbert et al., 2004), and trauma or cognitive arousal associated with the event (Dutton & Carroll, 2001). When recalling events, the perception of the event duration (Block, 1974), descriptions of people involved (Behrman & Davey, 2001), and the event's characteristics (Wells & Olson, 2003), have all been found to be susceptible to inaccuracies. It is also possible for a witness to be confident of a memory that has unknowingly been subject to external influences, irrespective of validation against other evidence. Ten months after the Flight 1862 accident, Crombag et al. (1996) found that 55 percent of participants reported witnessing the accident on TV, despite the fact that there was never any footage of the crash. Despite this, EWT is still perceived as a highly credible form of evidence compared to other evidence types (Shermer et al., 2011), meaning EWT can be strongly influential on the outcomes of the investigative process.

EWT in Air Accident Investigation

There is little research on EWT within the realm of air accident investigation. Aviation accidents and the resulting investigations differ highly from criminal contexts. Most notably, air accident investigations are conducted non-punitively, meaning the objectives of investigations are focused on making industry-wide safety recommendations as opposed to providing evidence to convict individuals via a criminal court. Accurate EWT can be used to verify events that occurred prior to the accident (Figure 1), allowing investigators to generate working hypotheses about the accident causation.

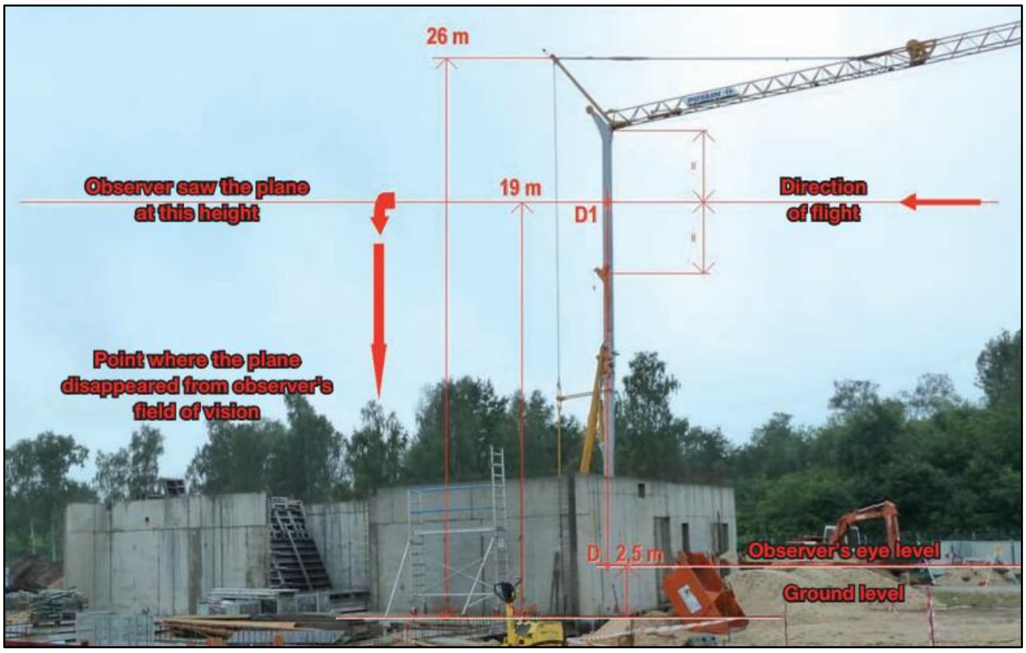

Figure 1. *EWT has been used to validate the movements of an aircraft before a crash by corroborating the testimony against landmarks in the environment (Frątczak & Konieczka, 2018, pg. 72).*

The Cognitive Interview (CI) is an EWT interview technique that has been found to considerably improve the accuracy of EWT (Memon et al., 2010). The CI is widely used in criminal and forensic investigations and has also been applied to accident investigation (Dodier et al., 2021). Regulatory and investigative bodies within the aviation industry appear to adhere to CI techniques as a means of obtaining EWTs following an accident. ICAO (n.d.) outlines guidance for investigators that recommends that EWTs are taken at the location of the accident, and highlights the importance of witnesses first being given the opportunity to describe their accounts of the events observed before being questioned by the investigator. The NTSB's (n.d.) procedure for obtaining EWTs follows this guidance, whilst also ratifying the importance of building rapport and avoiding leading questions.

However, there is evidence to suggest that EWTs in air accident investigations are still unreliable. English & Kuzel (2014) quantitatively analysed 239 witness reports of the Flight AA587 accident by positioning the location of each witness relative to the crash site and aircraft track. It was found there was considerable variation between witness accounts, and witness location was found to have no significant effect on the validity of the observations reported (Figure 2). Thus, the EWTs analysed were deemed to be highly unreliable.

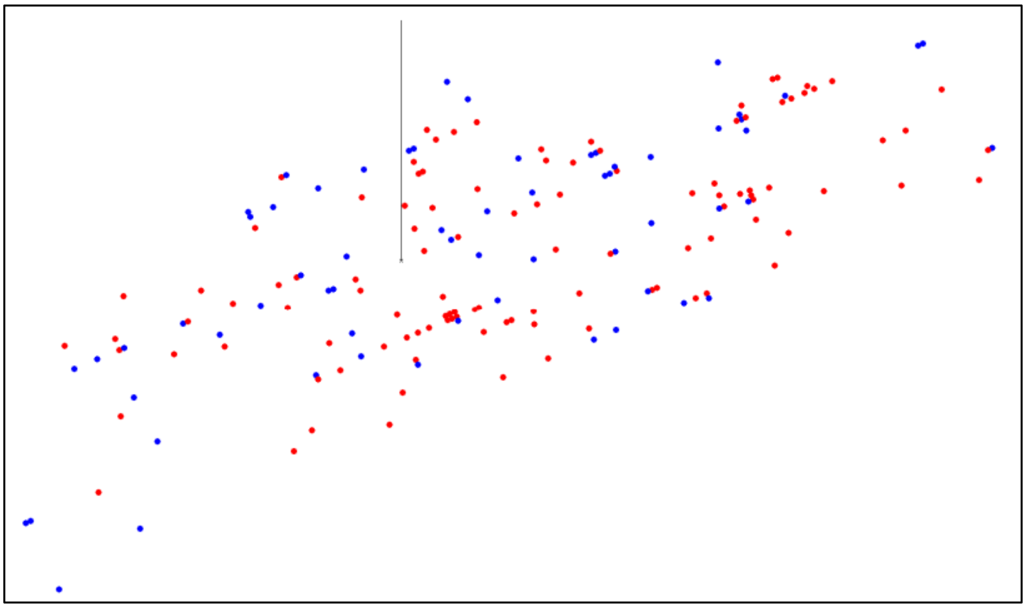

Figure 2. *Location of witnesses who observed the presence of fire (red) and no fire (blue) relative to the aircraft's flight path (line) (English & Kuzel, 2014, pg. 8).*

Solution: AI Application to EWT for Air Accidents

Artificial Intelligence (AI), the ability for a system to correctly interpret, learn from, use and adapt external data in relation to specific goals, is already being widely used in the aviation industry. Recently, EASA (2023) has proposed that AI can improve safety outcomes in aviation by helping infer knowledge through the understanding of large datasets. Despite this, new AI technologies are being woefully underutilised in the immediate investigative process following an air accident.

It is widely acknowledged that taking an EWT as soon as possible after the event improves testimony accuracy by reducing the influence of external factors (Wells et al., 2006). As such, the use of AI 'chatbots' to interview individuals immediately after witnessing a crime has been explored, with results suggesting that recall accuracy is improved (Minhas et al., 2022). The use of AI chatbots may improve recall accuracy by reducing perceived investigator bias and by following an interview schedule that contains appropriate questions, leading to a cost-effective and efficient solution to elicit accurate EWTs which are uninfluenced by external factors. It is thus proposed that the use of AI may be effective in improving EWT evidence quality in air accident investigations. AI chatbots are easier to administer than in-person interviews, meaning more EWT evidence can be collected within the accident environment in the immediate accident aftermath. An AI GPT with extensive knowledge on aviation accidents and CI techniques may enable eyewitnesses to better recall events in an un-biased setting after the accident using a rigorous, yet adaptive, questioning process. This process can also be used as a demonstrative tool to train investigators on effective questioning techniques in instances where an in-person interview is required. Investigators would have access to large eyewitness datasets soon after the accident, which could help validate other evidence sources during the investigation and assist in working hypotheses generation. Ultimately, an AI chatbot would not only save time and money, but would also provide investigators with high-quality validated evidence much earlier into the investigative process.

Proposed AI GPT Architecture

AI chatbots are typically based around generative pre-trained transformers (GPTs) that process natural language and provide output via conversation-based user interfaces, meaning they are intuitive to use and require little user training (Ray et al., 2023a). GPTs use a transformer architecture to understand and generate natural language. An AI GPT to assist in the air accident witness interview process should be primarily constructed using Large Language Model (LLM) methodology (Figure 3). LLMs seek to process and generate human-like text using neural networks that predict the probability of a word occurring in a given sequence which, when trained correctly, provide large quantities of linguistic knowledge and output (Petroni et al., 2019). Initially, the GPT should be pre-trained on large amounts of text data to gain a basic understanding of language generation. Following this, a fine-tuning phase consisting of a domain-specific LLM process would enable the GPT to gain extensive knowledge on the precursors to air accidents by training the model on large accident and incident report datasets. The LLM could also be trained on CI techniques during the fine-tuning phase, thus creating an interview protocol that accentuates accurate questioning. A chatbot user interface could then be linked to this LLM and, by using in-context learning, could ask questions that elicit the highest level of accuracy possible based on the information recalled. Questions asked by AI chatbot may be more ‘intelligent’ than those asked by an investigator as the GPT will have immediate access to a host of previous accident reports from which to work from. Once the GPT has effectively interviewed witnesses of the accident, it can use large vector databases to analyse the language used in the witness response via Natural Language Processing (NLP) functions and create a narrative of key findings, themes and events to assist investigators in generating a working hypothesis of the accident causation. For this output to be useful, investigators must be trained on developing prompting skills to ensure the GPT produces outputs that elicit the most useful data to aid the investigation.

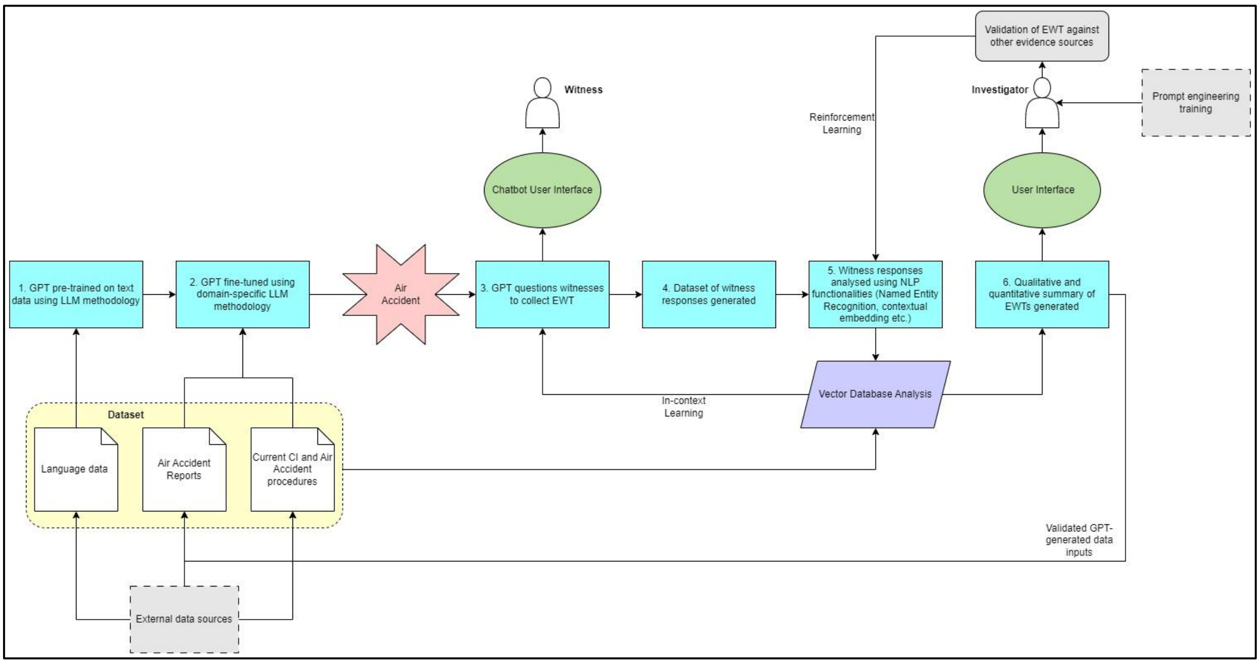

Figure 3. *Proposed architecture and process of the AI GPT.*

Concluding Remarks

Inaccurate EWTs have the potential to mar the quality of an investigative process, resulting in reduced safety outcomes and operational inefficiency for the investigating organisation. Within the realm of air accident investigation, EWT is diverse due to the complex nature of air accident events in a time-limited event window. The use of new AI GPT technology to assist in the collection and analysis of EWTs following an air accident aims to improve safety by improving the timeframe, efficiency and validation of witness information collected. This novel technique also ensures accurate EWT is collected, which can also be used to assist the investigation in validating other evidence types and help generate working hypotheses of accident causation. The AI chatbot interface enables large amounts of EWT to be collected soon after the accident occurs, meaning recall is less likely to be affected by external factors. The accessibility of this interface means investigator training centred around prompt engineering of the GPT can be conducted with minimal effort and resource consumption. It is evident that such technology has the potential to improve the quality of air accident investigations which may have far-reaching positive consequences for wider aviation safety in the future.

References

Behrman, B. W., & Davey, S. L. (2001). Eyewitness identification in actual criminal cases: An archival analysis. *Law and Human Behavior*, *25*(5), 475–491. https://doi.org/10.1023/a:1012840831846

Block, R. A. (1974). Memory and the experience of duration in retrospect. *Memory & Cognition*, *2*(1), 153–160. https://doi.org/10.3758/bf03197508

Crombag, H. F., Wagenaar, W. A., & Van Koppen, P. J. (1996). Crashing Memories and the problem of 'source monitoring.' *Applied Cognitive Psychology*, *10*(2), 95–104. https://doi.org/10.1002/(sici)1099-0720(199604)10:2<95::aid-acp366>3.0.co;2-#

Dodier, O., Ginet, M., Teissedre, F., Verkampt, F., & Fisher, R. P. (2021). Using the cognitive interview to recall real-world emotionally stressful experiences: Road accidents. *Applied Cognitive Psychology*, *35*(4), 1099–1105. https://doi.org/10.1002/acp.3828

Dutton, A., & Carroll, M. (2001). Eyewitness testimony: Effects of Source of Arousal on Memory, Source-monitoring, and Metamemory Judgments. *Australian Journal of Psychology*, *53*(2), 83–91. https://doi.org/10.1080/00049530108255128

EASA. (2023). *Artificial Intelligence Roadmap 2.0.* https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20#group-easa-downloads.

English, D., & Kuzel, M. (2014). Reliability of eyewitness reports to a major aviation accident. *International Journal of Aviation, Aeronautics, and Aerospace*. https://doi.org/10.15394/ijaaa.2014.1040

Frątczak, D., & Konieczka, R. (2018). Interviewing witnesses in proceedings related to an aircraft accident. *Issues of Forensic Science*, *300*(2), 67–75. https://doi.org/10.34836/pk.2018.300.3

Gabbert, F., Memon, A., Allan, K., & Wright, D. B. (2004). Say it to my face: Examining the effects of socially encountered misinformation. *Legal and Criminological Psychology*, *9*(2), 215–227. https://doi.org/10.1348/1355325041719428

ICAO. (n.d.). *Manual of Aircraft Accident and Incident Investigation, Part III Investigation* (9756-AN/965). https://skybrary.aero/sites/default/files/bookshelf/3708.pdf.

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of Automobile Destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, *13*(5), 585–589. https://doi.org/10.1016/s0022-5371(74)80011-3

Memon, A., Meissner, C. A., & Fraser, J. (2010). The cognitive interview: A Meta-analytic review and Study Space Analysis of the past 25 years. *Psychology, Public Policy, and Law*, *16*(4), 340–372. https://doi.org/10.1037/a0020518

Minhas, R., Elphick, C., & Shaw, J. (2021). Protecting victim and witness statement: Examining the effectiveness of a chatbot that uses artificial intelligence and a cognitive interview. *AI & SOCIETY*, *37*(1), 265–281. https://doi.org/10.1007/s00146-021-01165-5

NTSB. (n.d). Aviation Investigation Manual - Major Team Investigation. https://www.ntsb.gov/about/Documents/MajorInvestigationsManualApp.pdf.

Petroni, F., Rocktaschel, T., Lewis, P., Bakhtin, A., Wu, Y., Miller, A. H., & Riedel, S. (2019). Language Models as Knowledge Bases? *arXiv Preprint arXiv:1909.01066*.

Ray, A. T., Pinon-Fischer, O. J., Mavris, D. N., White, R. T., & Cole, B. F. (2023a). Aerobert-NER: Named-entity recognition for Aerospace Requirements Engineering using bert. *AIAA SCITECH 2023 Forum*. https://doi.org/10.2514/6.2023-2583

Shermer, L. O., Rose, K. C., & Hoffman, A. (2011). Perceptions and credibility: Understanding the nuances of eyewitness testimony. *Journal of Contemporary Criminal Justice*, *27*(2), 183–203. https://doi.org/10.1177/1043986211405886

Wells, G. L., & Olson, E. A. (2003). Eyewitness testimony. *Annual Review of Psychology*, *54*(1), 277–295. https://doi.org/10.1146/annurev.psych.54.101601.145028

Wells, G. L., Memon, A., & Penrod, S. D. (2006). Eyewitness evidence. *Psychological Science in the Public Interest*, *7*(2), 45–75. https://doi.org/10.1111/j.1529-1006.2006.00027.x