

University of Strathclyde

Genealogical, Palaeographic and Heraldic Studies

Centre for Lifelong Learning,

Faculty of Humanities and Social Sciences

How can Artificial Intelligence (AI) be used to empower genealogists in  
their research?

by

Cheryl Bollard

A dissertation presented in fulfilment of the requirements for the degree  
of MSc in Genealogical, Palaeographic and Heraldic Studies

2024

**Copyright statement:**

This dissertation is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

The copyright of this dissertation belongs to the author under the terms of the United Kingdom Copyright Acts as qualified by University of Strathclyde Regulation 3.50. Due acknowledgement must always be made of the use of any material contained in, or derived from, this dissertation.

Signed: Cheryl Bollard

Date: 17 June 2024

**Acknowledgements:**

I would like to thank the following people who assisted me during this research. My husband, Jeff Bollard, for his continued support for my study; as always, he went above and beyond to encourage me. My tutor, Andrew Douglas, who not only offered guidance and kept me on track but, just as importantly, shared the excitement of exploring a new technology. The Council and members of Family History ACT, my local genealogy society, where I was first introduced to generative-AI, in particular the volunteers who participated in the research. Finally, I'd also like to thank Steve Little and the National Genealogical Society; the journey would have been far more difficult without their courses and support.

## Abstract

AI is increasingly integral to our lives, reshaping how we work, think and interact. In genealogy, AI, particularly generative-AI, has significant potential to transform research practices. This study explores how AI can empower genealogists by streamlining research tasks and uncovering new insights into historical data. A mixed-methods evaluation involved genealogists testing generative-AI in workflows and interviews on three AI-enhanced historical projects. The findings indicate that integrating AI into genealogy workflows augments genealogists' skills and knowledge, improving their efficiency and effectiveness. However, it also highlights the need for caution to maintain the accuracy and reliability of genealogical research. The study recommends updating standards and developing comprehensive guidelines for ethical AI use. Structured educational programs are advised to equip genealogists with a thorough understanding of AI's benefits and limitations, ensuring responsible integration into their work.

**Keywords:** generative AI, genealogy workflows, ethical AI use, historical data analysis, AI education in genealogy

## Contents

Abstract.....	3
Table of Figures.....	6
Table of Tables.....	8
Glossary/ Nomenclature.....	9
Terminology used in this dissertation.....	11
Chapter 1. Introduction.....	12
Chapter 2 – Literature Review.....	18
2.1 AI 101 - AI Concepts.....	18
2.2 Impact of AI on Traditional Genealogical Methods.....	22
2.3 Adaption and Change Management in Genealogy.....	25
2.4 Role of AI in Genetic Genealogy.....	27
2.5 With Great Power comes Great Responsibility – Ethically using AI in Genealogy.....	30
2.6 Literature Review Conclusion.....	32
Chapter 3 – Research Methods.....	33
3.1 AI’s Capabilities in Genealogy (Research Questions 1-5).....	34
3.2 The Impact of AI Integration on Genealogists: RQ 6.....	43
3.3 Limitations and Mitigation Strategies.....	44
3.4 Methodology Conclusion.....	45
Chapter 4 – Empirical Findings: Description, Analysis & Synthesis.....	46
4.1 Generative-AI’s Capabilities in Genealogy.....	46
4.2 The Ethical Considerations of AI in Genealogy.....	71
4.3 The Impact of AI Integration on Genealogy:.....	79
4.4 Future Prospects and Practical Implications of AI Integration in Genealogy:.....	108
4.5 Conclusion.....	110
Chapter 5 – Conclusion.....	111
5.1 Generative-AI’s Capabilities in Genealogy.....	111
5.2 The Ethical Considerations of AI in Genealogy.....	112
5.3 The Broad Impact of AI on Genealogy.....	113
5.4 Contribution to Knowledge.....	114
5.5 Self-reflection.....	115
5.6 Future Directions.....	116
5.7 Conclusion.....	117
Appendix.A Capability Analyses.....	<b>Error! Bookmark not defined.</b>
A.1 Preliminary Questionnaire.....	<b>Error! Bookmark not defined.</b>
A.2 Capability Analyses - General Instructions.....	<b>Error! Bookmark not defined.</b>
A.3 Capability Analysis Test 1.....	<b>Error! Bookmark not defined.</b>
A.4 Capability Analysis Test 2 – Extraction of information..	<b>Error! Bookmark not defined.</b>
A.5 Capability Analysis Test 3 – Translation.....	<b>Error! Bookmark not defined.</b>

A.6 Capability Analysis Test 4 - Data Interpretation.....	<b>Error! Bookmark not defined.</b>
A.7 Capability Analysis Test 5 – Complex Task .....	<b>Error! Bookmark not defined.</b>
A.8 Capability Analysis Test 6 – Ethics .....	<b>Error! Bookmark not defined.</b>
A.9 Capability Analysis Test 7 – DNA and Genetic Genealogy.....	<b>Error! Bookmark not defined.</b>
A.10 Capability Analysis Test 8 - Document Search.....	<b>Error! Bookmark not defined.</b>
A.11 Debriefing Questionnaire .....	<b>Error! Bookmark not defined.</b>
Appendix.B Raw Data .....	<b>Error! Bookmark not defined.</b>
Appendix.C Tabulated Data .....	<b>Error! Bookmark not defined.</b>
Appendix.D Example of Prompt Engineering .....	<b>Error! Bookmark not defined.</b>
D.1 What is Prompt Engineering? .....	<b>Error! Bookmark not defined.</b>
D.2 It's All in the Prompt.....	<b>Error! Bookmark not defined.</b>
Appendix.E AI generated Ethical Framework .....	<b>Error! Bookmark not defined.</b>
Appendix.F Project Interview Reports .....	<b>Error! Bookmark not defined.</b>
F.1 Project 1 Interview 1 Ironclad Sisterhood – Andrew Redfern.....	<b>Error! Bookmark not defined.</b>
F.2 Project 2 Interview: Landscape of Production & Punishment – Richard Tuffin.....	<b>Error! Bookmark not defined.</b>
F.3 Project 3 Interview: Making Crime Pay – Professor Hamish Maxwell-Stewart....	<b>Error! Bookmark not defined.</b>
Appendix.G The Tasmanian Longitudinal Dataset .....	<b>Error! Bookmark not defined.</b>
Appendix.H Generative AI Prompts and Outputs .....	<b>Error! Bookmark not defined.</b>
End Notes .....	119
Bibliography .....	131
Primary Sources.....	131
Secondary Sources.....	132

## Table of Figures

Figure 1 Milestones in OpenAI’s Generative AI capabilities illustrating the speed of change. ....	12
Figure 2 Comparison of generative-AI models June 2024. ....	16
Figure 3 Venn diagram depicting the relationship between AI, ML, and Deep Learning. ....	19
Figure 4 Use of human feedback to improve the performance of GPTs.....	20
Figure 5 Dell’Acqua’s “Jagged frontier of AI capabilities”. ....	21
Figure 6 The National Library of Australia’s shift from a card catalogue to a computer system. ....	23
Figure 7 Genealogical research methods timeline. ....	25
Figure 8 Rogers’ Theory of Diffusion of Innovation. ....	25
Figure 9 Use of AI technologies in genetic genealogy. ....	27
Figure 10 Example of Capability Analyses. ....	36
Figure 11 Genealogical and computer skill level of the testing cohort.....	39
Figure 12 Age of testing cohort. ....	40
Figure 13 Example of steps for inductive coding of quantitative data.....	41
Figure 14 The testing cohort’s diverse genealogical objectives. ....	48
Figure 15 Evaluation of CA5.2 summarisation task.....	49
Figure 16 AI-generated summary of Coe ARNOLD’s will. (Transcript in Appendix A CA5) ....	50
Figure 17 Comparison of the time taken for generative-AI and manual summary....	51
Figure 18 Evaluation of extraction accuracy.....	52
Figure 19 Excerpts illustrating AI extraction of information from Coe ARNOLD’s will. (Transcript in Appendix A CA5). ....	54
Figure 20 Evaluation of translation tasks.....	55
Figure 21 Comparison of FamilySearch manual translation with generative-AI translation. ....	57
Figure 22 Example of AI-generated narrative from an obituary – CA2.3. ....	58
Figure 23 Evaluation of generation tasks. ....	59
Figure 24 Preliminary Survey results on prior use of AI for genealogy workflows. ...	60
Figure 25 Quantitative analysis of generative AI’s performance across all task categories.....	62
Figure 26 Analysis of the results for CA6.3.2 illustrating the high variability of responses.....	65
Figure 27 Box and Whisker chart showing the distribution of scores for CA5.3.2. ...	66
Figure 28 Analysis of generative-AI’s performance in “jagged frontier” tasks.....	68
Figure 29 Illustration of generative-AI ability to create a family tree. ....	69
Figure 30 Ethical concerns identified by testing cohort. ....	71
Figure 31 Evaluation of respondents’ concern on the use of AI-assisted genealogy. ....	72
Figure 32 Application of Rogers’ Innovation-Decision Process. ....	72
Figure 33 Evaluation of applicability of existing genealogical standards. ....	73
Figure 34 Ethical integration framework suggested by testing cohort. ....	75
Figure 35 Respondents’ recommendations for implementing ethical AI-assisted research. ....	76
Figure 36 Evaluation of generative-AI as an assistant in resolving ethical questions. ....	77
Figure 37 The prevalence of DNA testing amongst respondents. ....	79

Figure 38 Results of integration of genetic and traditional genealogy methods. ....	80
Figure 39 Frequency and usage of DNA AI tools. ....	81
Figure 40 Results showing benefits of DNA testing. ....	83
Figure 41 Evaluation of concerns about the use of AI in genetic genealogy. ....	85
Figure 42 The stages of converting handwriting to text using consumer-level AI. ....	87
Figure 43 Example of a MyHeritage Natural Language search. ....	89
Figure 44 Evaluation of MyHeritage AI Record Finder. ....	90
Figure 45 Search strategies used by respondents to find a UK birth record. ....	92
Figure 46 Example of an AI-generated research plan to locate a UK birth record (CA8.2) ....	93
Figure 47 Evaluation of AI-generated research plans. ....	94
Figure 48 Evaluation of bias in online research. ....	96
Figure 49 "Beautification" of convict and acknowledgement of AI-generated images. .....	98
Figure 50 LiDAR-generated elevation map of Port Arthur showing significant activities. ....	100
Figure 51 The Port Arthur Convict Coal Mines showing terrain visible from ground level ....	101
Figure 52 LiDAR scan of the quarry site revealed a spoil heap and tramlines. ....	101
Figure 53 The process from convict conduct record to table and GIS geo-located map. ....	102
Figure 54 The Convict Memorial "Unshackled" at Hobart Penitentiary Chapel. ....	103
Figure 55 Close-up of one side of the memorial pillar in the Unshackled immersive experience. ....	104
Figure 56 Convict record of Edward GRIFFITHS and digital avatar made from description. Note the inclusion of tattoos. ....	105
Figure 57 Evaluation of AI's ability to empower genealogists. ....	108
Figure 58 Distribution of scores for AI empowerment of genealogy workflows. ....	109
Figure 59 Example of CA6 results from Respondent #3. ....	<b>Error! Bookmark not defined.</b>
Figure 60 Example of tabulated results. ....	<b>Error! Bookmark not defined.</b>
Figure 61 Example of narrative created with a basic prompt. ....	<b>Error! Bookmark not defined.</b>
Figure 62 Example of a narrative created with a detailed prompt. ...	<b>Error! Bookmark not defined.</b>
Figure 63 Example of a narrative generated with a persona prompt. ....	<b>Error! Bookmark not defined.</b>

## Table of Tables.

Table 1 Correlation between the theoretical framework and the research questions.	33
Table 2 Structure and content of the Capability Analyses.	35
Table 3 Mapping of research questions to the theoretical framework.	46
Table 4 Use of Generative-AI platforms for CA evaluation.	47
Table 5 Overview of generative-AI tasks by category.	48
Table 6 Comparison times for CA5.2.3 - Extraction	53
Table 7 Sample of qualitative feedback illustrating the positive impact of generative AI.	61
Table 8 Relationship between accuracy, efficiency and usefulness.	63
Table 9 Sample of qualitative feedback illustrating some of generative-AI's limitations.	64
Table 10 A sample of qualitative responses illustrating the tedious nature of verification.	65
Table 11 Beginner-level scores for CA5.3.2.	67
Table 12 Outline of AI-generated framework for ethical AI-assisted research.	74
Table 13 Use of ethnic admixture results	84
Table 14 Heatmap showing respondents' options if a record was unavailable online.	95
Table 15 Overview of AI-enhanced projects to reinterpret historical convict records.	97
Table 16 Example of TLD datasets.	99
Table 17 Predictions for the use of AI in genealogy in 2030.	110
Table 18 List of documents in Appendix A	<b>Error! Bookmark not defined.</b>
Table 19 Example of Little's prompt engineering approach.	<b>Error! Bookmark not defined.</b>
Table 20 List of project interview reports.	<b>Error! Bookmark not defined.</b>
Table 21 Tasmanian Longitudinal Dataset record sources.	<b>Error! Bookmark not defined.</b>



## Glossary/ Nomenclature

<b>Abbreviation</b>	<b>Explanation</b>
AI (Artificial Intelligence)	The broad concept of machines performing tasks that typically require human intelligence.
AGGRA	The Australasian Association of Genealogists and Record Agents is a professional association for genealogists.
BCG	Board for Certification of Genealogists – a US-based professional association for genealogists
Deep Learning	A specialisation within ML involving neural networks that mimic human brain functioning, useful for complex tasks like image and speech recognition.
DNA Genotyping	The process of examining an individual's DNA at specific locations, often for identifying ethnic admixture or haplogroup details.
Ethnic Admixture	A term describing the mixing of ancestries or ethnic groups in an individual's genetic makeup.
Family History ACT FHACTION	Family History ACT is a genealogical society in Canberra and the Australian Capital Territory.
GDPR (General Data Protection Regulation)	A comprehensive data protection law that became enforceable in the European Union (EU) on May 25, 2018. It's designed to give individuals control over their personal data and simplify international business's regulatory environment by unifying the regulation within the EU.
Generative AI (Generative-AI)	This term refers to a form of AI designed to create new content rather than analyse or process existing data. It includes GPTS such as ChatGPT and Copilot and image-to-text generators.
Generative Pre-trained Transformers (GPT)	A type of Large Language Model known for its ability to generate coherent text. Examples include ChatGPT, Gemini, Microsoft Copilot and Claude.
Genome (Human Genome)	The complete set of DNA in an organism, including all of its genes. In humans, the genome consists of 23 chromosome pairs.
Global Information Systems (GIS)	A framework for gathering, managing, and analysing geographical data. It analyses spatial location and organises layers of information into visualisations using maps and 3D scenes, providing deeper insights into data, such as patterns, relationships, and situations.
Haplogroup	A group of similar haplotypes that share a common ancestor with a single SNP mutation.
Haplotype	A DNA sequence at adjacent locations on a chromosome that are inherited together. Haplotypes can be used to identify genealogical lineage and genetic variations.
IEEE	Institute of Electrical and Electronic Engineers. A professional association dedicated to advancing technology for the benefit of humanity.
In Common With (ICW)	It is used in genetic genealogy to describe matches that two individuals share with the same third part. (shared matches)
ISOGG	The International Society of Genetic Genealogy is an organisation dedicated to promoting the use of genetics in genealogical research.

Large Language Models (LLM)	A specific type of Natural Language Processing application, trained on extensive text data, is capable of generating, understanding, and responding to human language – sometimes called a chatbot.
LiDAR Light Detection and ranging Aerial Remote sensing imagery	LiDAR is a remote sensing method that uses light in the form of a pulsed laser to measure variable distances to the Earth. These light pulses—combined with other data recorded by the airborne system—generate precise, three-dimensional information about the shape of the Earth and its surface characteristics. LiDAR technology is widely used for making high-resolution maps, with applications in geology, seismology, forestry, and urban planning, among others.
Machine Learning (ML)	A subset of AI where systems learn and make predictions from data.
MRCA Most Recent Common Ancestor	In genetic genealogy, MRCA refers to the most recent individual from whom two or more shared matches are directly descended. It helps identify a point in ancestral lines where their paths converge.
mtDNA	The DNA found in mitochondria, inherited from the mother, is used in studies tracing maternal lineage.
Natural Language Processing (NLP)	The application of AI to understand and interpret human language.
Neural Network	A type of Deep Learning that mimics the way neurons in the human brain connect.
Next Generation Sequencing	A method of DNA sequencing that allows for the rapid sequencing of the entire genome or parts of it, significantly faster and cheaper than traditional methods.
Non-Paternity Event (NPE)	Also known as Not Parent Expected – the person recognised as the parent is not the biological parent.
RQG	Register of Qualified Genealogists, a professional association for genealogists
SNP (Single Nucleotide Polymorphisms)	Genetic variations at specific positions in the genome, crucial for genetic analysis.
WATO	What are the Odds - a tool used to explore and hypothesise relationships between DNA matches based on probability.
Y-DNA	The DNA found on the Y chromosome, inherited paternally, is used in studies tracing paternal lineage and male ancestry.

## Terminology used in this dissertation

Although the term “AI” is often applied to mean the newest and most human-like forms of artificial intelligence, this dissertation distinguishes between several different forms of AI. “Generative-AI” refers specifically to applications such as ChatGPT or DALL-E 2, which are designed to create new content. At the same time, Large Language Model (LLM) refers to generative-AI chatbots such as ChatGPT, Claude and Copilot. “AI” is used to refer to other artificial intelligence technologies.

## Acknowledgement of the use of Generative-AI

I would like to acknowledge the use of generative-AI in this dissertation. Since this research focused on AI’s ability to empower genealogists in their work, the exploration and application of generative AI was a core aspect of the dissertation. When generative AI content was used, it was properly cited and clearly identified as AI-generated in the references.

ChatGPT-4 was used as a collaborative tool to assist with the initial planning and early revisions of my work in much the same way as a friend or colleague. It was used to suggest potential resources for further research and assist in developing new skills, including using Python for statistical analysis. While ChatGPT-4 served as a tool to facilitate the research process, the formulation of the research questions, the interpretation of data, the synthesis of information, and the critical analysis are entirely my own work. The final content, strategic direction, and conclusions reflect my personal insights and academic judgment.

# Chapter 1. Introduction

## 1.1 Background

Artificial Intelligence (AI) has been widely recognised for its potential to improve workflows across many fields, including genealogy, where it has helped reshape genealogical research methodologies over many years by expediting DNA analysis, processing large datasets and improving information retrieval processes.<sup>1 2 3</sup> In late November 2022, the AI world was transformed with the release of ChatGPT, a generative-AI chatbot that was affordable, user-friendly and accessible.<sup>4</sup> With over a million users within five days and ten million within a few months, it seemed like AI exploded overnight from an obscure academic study to a powerful force transforming every aspect of life.<sup>5</sup> Mollick compares AI to other General-Purpose Technologies that have transformed everyday life, like electricity, computers and the internet, but notes its impact has been felt in months rather than decades.<sup>6</sup> (Fig.1)

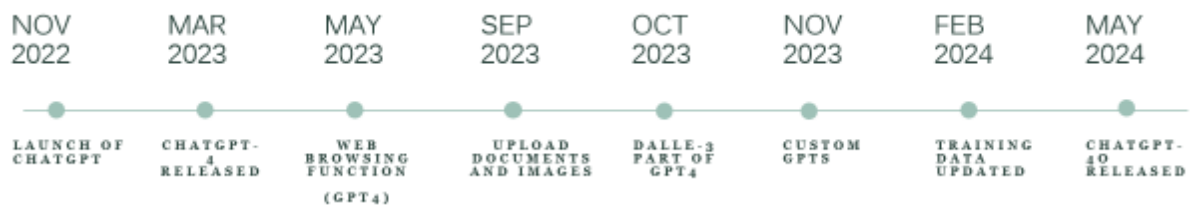


Figure 1 Milestones in OpenAI's Generative AI capabilities illustrating the speed of change.<sup>7</sup>

Generative-AI creates new content, unlike other forms of AI that analyse or process existing data. Models like ChatGPT, Co-Pilot, Claude or Midjourney generate text, images and other media based on patterns learned from large amounts of training data.<sup>8</sup> Studies show that knowledge workers using generative-AI improved productivity by 20-80%.<sup>9 10</sup> When combined with its ability to synthesise text, predict data patterns, and automate complex tasks, generative-AI has the potential to provide genealogists with new ways to enhance their research.

However, early reviews and feedback from genealogists did not entirely support these optimistic findings. Initial reports expressed scepticism and reluctance within the genealogical community, with genealogists expressing concerns about the reliability and accuracy of AI-generated information.<sup>11 12 13</sup> Generative-AI's tendency to create non-factual information, known as hallucinations, poses significant problems in a field that relies heavily on accuracy and verifiable data, and many questioned the value of generative-AI in genealogy.<sup>14 15 16</sup>

However, as the technology evolved, genealogists continued experimenting and recognised more of generative-AI's potential and adopted it into their workflows. Bettinger, Little and Mollick demonstrated that clear prompts and contextual information significantly enhanced generative-AI's performance, decreasing hallucinations, reducing the risk of errors and improving the overall quality of genealogical research.<sup>17 18 19</sup> They proposed that the best approach to integrating generative-AI into genealogy was to augment genealogists' skills through collaboration.<sup>20</sup>

By working collaboratively with AI, routine or time-consuming tasks such as data extraction or summarisation can be streamlined, allowing genealogists to focus on more complex analytical tasks requiring human judgment and contextual understanding.<sup>21</sup> This partnership not only speeds up the research process but also enhances the quality and depth of the analysis, allowing genealogists to make more informed decisions and explore new research avenues.

Given these insights, examining how genealogists can work collaboratively with generative-AI in traditional genealogy workflows is important. The rapid evolution of these technologies and the current gaps in practical knowledge call for careful examination of how AI can be integrated into traditional genealogy workflows while maintaining accuracy and innovation.

## **1.2 Research Focus**

Generative-AI is sometimes seen as a magic wand, instantly solving research questions or miraculously finding long-missing records to solve genealogical brick walls. In reality, generative-AI empowers research in small increments by streamlining tasks, freeing time for further research.<sup>22 23</sup> Its practical application in genealogy requires realistic expectations and an understanding of its capabilities, limitations and ethical implications like data privacy, accuracy and historical integrity.

This research aimed to critically assess the integration of AI into genealogy workflows by investigating its actual performance when undertaking routine genealogical tasks like data extraction, document analysis and translation. Focusing on these practical applications, the research sought a balanced view of AI's potential, moving beyond media claims to evaluate its real-world effectiveness.

This research also provided an opportunity to explore the initial impact of this emerging technology (generative-AI) in genealogy. The last significant studies on technology adoption within genealogy were conducted in 2016 with the move towards digital technologies and in 2020 concerning the use of DNA in genealogical practices.<sup>24 25</sup> As generative-AI continues to evolve and more genealogists experiment and adopt this technology, genealogical research methods are likely to change considerably. This makes the current study both timely and essential, providing a foundation for understanding how genealogists are adapting to this transformative tool at an early stage.

### **1.3 Overall Research Aim and Individual Research Objectives.**

The primary aim of this research was to investigate the potential of AI to empower genealogists by exploring the benefits and limitations of its ethical integration into genealogy workflows. In this context, empowering means providing the tools, skills or knowledge to do their research and other related work more effectively.

The specific objectives of this research are:

1. What are the benefits and limitations of generative-AI for genealogy purposes?
2. How can generative-AI be used in genealogy, how effective, efficient, and accurate is it for genealogical purposes, and how aware are genealogists of these tools?
3. What ethical considerations and safeguards are needed when using AI, and what risks arise from not implementing them?
4. How is AI being integrated into current genetic genealogy processes, and what are the benefits and challenges?
5. How is the use of AI changing the availability, access, and use of historical documents?
6. How is AI being used to re-interpret historical data to provide new insights for genealogical research?

One of the challenges of integrating AI into genealogy is the newness of the technology. AI is still being explored and tested to understand its capabilities, limitations, and the best way to utilise it.<sup>26</sup> These research objectives provide a systematic framework for examining and evaluating AI's use in key areas of genealogical interest, addressing whether AI can empower genealogists in their work.

## **1.4 Research methods**

This research used a literature review and empirical data to investigate the emerging use of AI and generative-AI in genealogy. Given the recent emergence of generative-AI, traditional academic publications are scarce and were supplemented with material from online forums, including Facebook and blogs.

Empirical data for Research Questions.1-5 was gathered through benchmark testing (Capability Analyses (CA)), using a mixed methods survey to evaluate generative-AI's performance in routine genealogical tasks. This approach, recommended by Cresswell, combines statistical information and qualitative feedback for an in-depth analysis of complex issues.<sup>27</sup> Similar methodologies were employed in studies examining generative-AI use in professional settings.<sup>28 29</sup> The CAs were completed by genealogists from Family History ACT (FHACTION). This group is representative of many genealogists who, although not paid to research, take a professional approach to their research.

Data for Research Question.6 was collected in interviews with three project leaders involved in projects using AI to reinterpret historical data. The methodology aligned with the principles outlined by Yin and Denscombe and provided an in-depth exploration of the integration of AI in a real-life context.<sup>30 31</sup>

The combination of the literature review and empirical research allowed for the comparison of theory and practice and provided a comprehensive analysis of the applicability of AI integration into genealogy. Full details of the research methods used for the empirical study, including justification for the chosen methods, can be found in Chapter 3.

## **1.5 Choice of Generative-AI Model**

This research required an AI model with advanced functionality, high accuracy and versatility to fully explore generative-AI's capabilities in genealogy, features usually found in cutting-edge models.<sup>32</sup> For most of the research period, ChatGPT-4 was used as it was the most advanced generative-AI model, offering features like document upload, real-time internet access, and custom GPTs.<sup>33</sup> These capabilities significantly reduced AI hallucinations and produced more accurate and contextually appropriate responses. However, other generative-AI models were used for comparison. By May/June 2024, technology had evolved with frontier models like GPT-4o, Gemini 1.5, and Claude 3 Opus, with improved

skills in reasoning over large amounts of text, generating ideas and writing non-robotically and was used for exploration during the latter part of my research.<sup>34</sup> (Fig.2)

Model	Connected to the Internet	Makes images	Runs code	Does data analysis	Sees images	Sees video	Reads files	Works with your documents	GPTs	Superpower
GPT-4o	Yes	Yes	Yes	Yes	Yes	No	Yes	No	Yes	The best general model for doing stuff. Multimodal voice is coming and will be a big deal
Copilot for web (using GPT-4)	Yes	Yes	No	No	Yes	No	Yes	Yes	Coming	Part of the suite of Copilots, works directly with MS Office and Windows in interesting ways
Claude 3 Opus	No	No	No	No	Yes	No	Some	No	No	Excellent writer, high levels of "insight"
Gemini 1.5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Connects seamlessly to Google utilities. Huge context window

Figure 2 Comparison of generative-AI models June 2024. <sup>35</sup>

Each testing cohort member chose the model they believed was best suited for their specific tasks. This varied approach allowed them to leverage the strengths of different generative-AI models, ensuring robust and reliable research results, and is discussed further in Chapter 4.

### 1.6 Research Value.

This research benefits the genealogy community by providing insights into integrating AI and generative-AI into genealogy workflows, filling gaps in our understanding of AI's potential impacts and practical applications. It is one of the earliest formal studies that empirically evaluate the capabilities and limitations of AI in genealogy, offering a balanced perspective that challenges early misconceptions of AI's usefulness and tests theoretical propositions on AI's applicability in a field based on factual accuracy and reliability.

The study demonstrated AI's impact on genealogy research methods, highlighting the benefits and limitations of generative-AI for genealogy. The findings supported the literature and demonstrated that generative-AI, when used collaboratively, augments genealogists' performance and empowers their workflow. The study also contributed to broader discussions on the ethical use of AI and the need for robust standards and guidelines to safeguard against potential risks while maximising the benefits.



This research contributes to the bank of knowledge and skills genealogists need to integrate AI into their research effectively, helping to keep methodologies current and relevant. The involvement of a local grassroots genealogy society in this research fosters a community of practice that supports the informed dissemination of AI knowledge, preparing genealogists for increasingly tech-driven research. This collaborative aspect provides a replicable model for other genealogy societies and highlights the communal benefits of such research.

The next chapter examines the literature relevant to the objectives of this research, beginning with an introduction to AI.

## Chapter 2 – Literature Review

This research used a literature review and empirical data to investigate the emerging use of AI and generative-AI in genealogy. Given the recent emergence of generative-AI, traditional academic publications such as journals, theses and books are just becoming available. Existing literature is primarily from practitioners, highlighting a significant gap in academic research into AI's role in genealogy. Video presentations, real-time discussions, and experiences shared in Facebook groups and blogs supplemented traditional sources. These digital forums provided valuable insights into current practices for integrating AI into genealogical workflows, filling gaps in the academic literature.

The review starts by investigating the use of AI in genealogical research and how genealogists adapt and manage change. It also analyses the role of AI in genetic genealogy and the ethical implications associated with using AI. These themes address fundamental questions about how and when genealogists use AI tools and provide a framework for the empirical investigation's design and analysis.

### **2.1 AI 101 - AI Concepts**

This section begins with an overview of AI technologies before discussing their impact on genealogy and highlighting some AI innovations in the field, setting the stage for understanding AI's technological implications in genealogical research.

Artificial Intelligence, Machine Learning (ML), and Deep Learning (DL) are often used interchangeably but represent different layers of complexity and capability in data processing and decision-making. Goodfellow describes AI as the broadest concept, encompassing tasks that typically require human intelligence, such as visual perception.<sup>36</sup> Machine Learning, a subset of AI, uses algorithms to analyse data, learn from it and make predictions about complex patterns.<sup>37</sup> Deep Learning, a further specialisation within ML, involves neural networks with multiple layers, making it effective at handling large volumes of data and complex tasks like image and speech recognition.<sup>38</sup> Goodfellow asserts Deep Learning is the current cutting-edge of AI's capacity to mimic human cognitive functions.<sup>39</sup> (Fig.3)

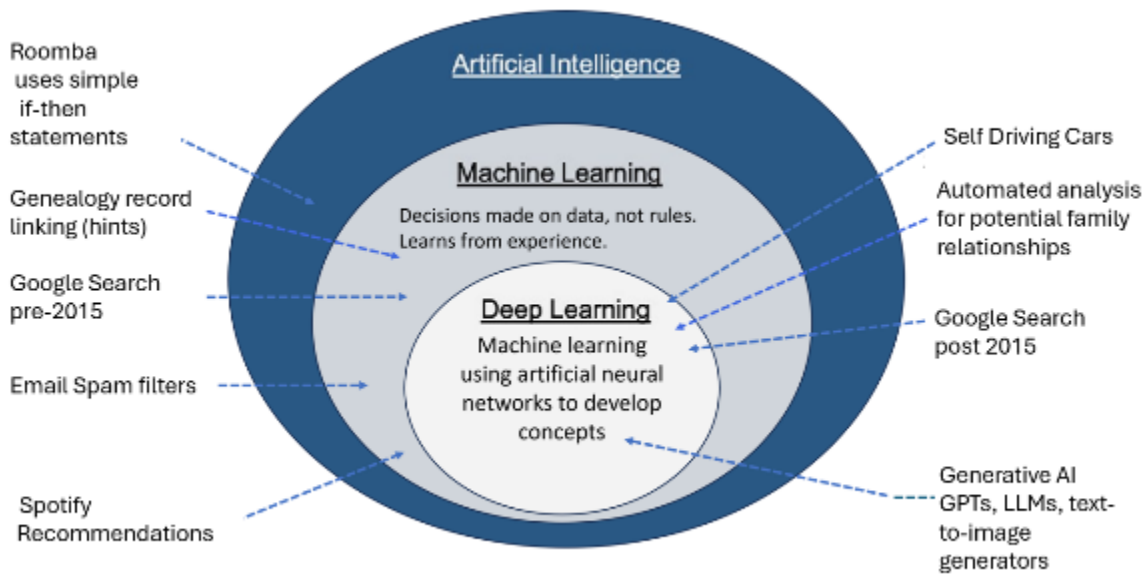


Figure 3 Venn diagram depicting the relationship between AI, ML, and Deep Learning.<sup>40</sup>

Pillai's analysis puts this into a genealogical context, explaining AI provides the data analysis framework, ML offers strategies to gather insights from historical records, and Deep Learning enables the detailed analysis of complex genealogies to suggest potential family connections.<sup>41</sup>

### **Generative AI - Reshaping how Genealogists Work**

Generative-AI is revolutionising how genealogists and other professionals work, by transforming traditional methods with advanced natural language processing (NLP). NLP includes DL applications like language translation and sentiment analysis to process and understand human language.<sup>42</sup>

A form of generative AI, Large Language Models like ChatGPT are an extension of NLP trained on extensive datasets to generate, understand, and respond to human language.<sup>43</sup> LLMs can help with a wide range of intellectual tasks, and as Mollick claims, working with them is like working with an intelligent person who sometimes makes mistakes; they can do amazing things but are frustrating at times.<sup>44</sup>

The generative-AI training process involves unsupervised learning and reinforcement learning from human feedback.<sup>45</sup> Initially, LLMs undergo unsupervised learning, analysing massive amounts of text data to identify patterns and understand language structures.<sup>46</sup> This pre-training enables the system to generate text in a stochastic manner, using probabilistic algorithms to produce coherent and contextually relevant text based on its input.<sup>47</sup> Reinforcement learning from human feedback fine-tunes the model's responses to align more closely with human values and preferences, enhancing the model's accuracy, reliability, and adherence to ethical guidelines by reducing biases and inappropriate content.<sup>48</sup> (Fig.4) Hörnemalm emphasises that the training data's quality, variety and size determine generative-AI's accuracy and biases.<sup>49</sup>

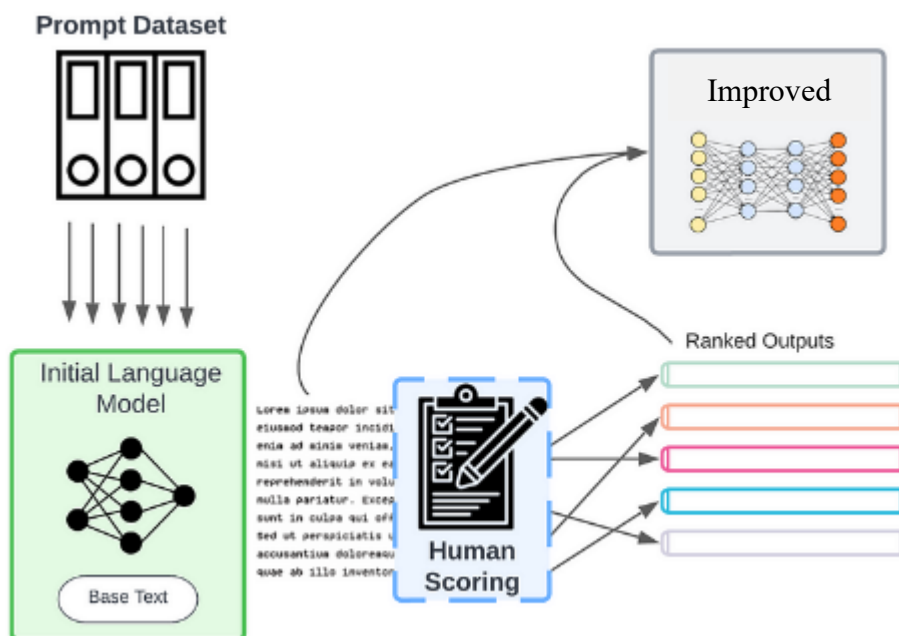


Figure 4 Use of human feedback to improve the performance of GPTs.<sup>50</sup>

One of the significant challenges with generative-AI is hallucinations, where models generate incorrect or fabricated information. This occurs when the LLM is given ambiguous context and uses its non-deterministic nature and probabilistic sampling to produce a plausible yet inaccurate response.<sup>51 52</sup> Kellogg and Mollick recommend prompt engineering as a valuable strategy to mitigate this issue.<sup>53 54</sup> Prompt engineering involves fine-tuning AI models through carefully crafted input prompts to enhance response accuracy and contextual relevance.<sup>55</sup> Little and Mollock use techniques like crafting clear and specific prompts, providing contextual information to minimise misinterpretations, iterative refinement to continuously improve prompts based on AI outputs (prompt chaining), and anchoring

responses to well-established facts.<sup>56 57</sup> Their research demonstrates that these strategies can significantly enhance the reliability of AI-generated content, proving particularly beneficial in fields demanding high accuracy and trustworthiness, including genealogy.

Research by Dell’Acqua highlighted the benefits of LLMs in enhancing work efficiency and decision-making of knowledge workers like genealogists.<sup>58</sup> Their study demonstrated that consultants working collaboratively with ChatGPT-4 significantly outperformed those who didn't use AI across various tasks.<sup>59</sup> The results showed improvements in performance speed, quantity and quality, demonstrating the substantial impact of generative-AI in empowering workers’ capabilities.<sup>60 61</sup> The research identified two successful patterns of AI integration: the "Centaur," who strategically divide tasks between AI and human effort, and the "Cyborgs," who integrate AI seamlessly into their workflow.<sup>62</sup>

Although LLMs are good at many tasks, there are some they do poorly and others not at all. Dell’Acqua’s "jagged frontier" concept illustrates the uneven capabilities of AI across tasks of similar difficulty and highlights the importance of being aware generative-AI’s limitations.<sup>63</sup> The research demonstrated that generative-AI performs well for the tasks inside the frontier, but the probability of hallucinations increases for tasks on the frontier. (Fig.5)



Figure 5 Dell'Acqua's "Jagged frontier of AI capabilities".<sup>64</sup>

This understanding of AI's strengths and limitations offers valuable insights for genealogists and professionals, helping them to integrate AI into their work processes while maintaining an awareness of over-reliance on the technology.<sup>65</sup>

Dell'Acqua's, Little and Mollick's research played an essential role in the design and methodology of my research's data collection.

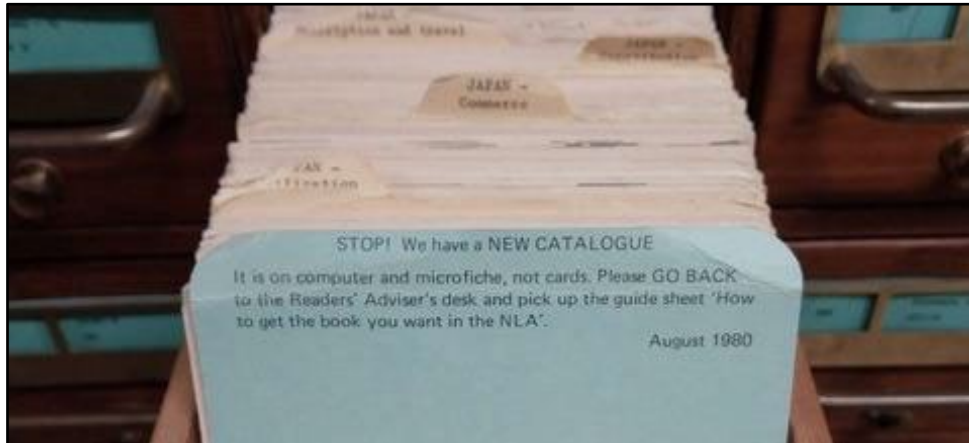
## **2.2 Impact of AI on Traditional Genealogical Methods**

As Pillai noted, AI has transformed genealogical research methods by enhancing search optimisation, recommendation systems and document management systems for faster digitisation, indexing and processing of records.<sup>66 67</sup> Its integration into genealogy is evident in family tree hints produced by sophisticated AI systems transcribing historical documents, linking records, and mapping family relationships.<sup>68</sup>

New genealogical data collections are becoming available as AI handwriting text recognition algorithms (HTR) accelerate their accessibility.<sup>69</sup> HTR technology transcribed the 1950 US Census in nine days compared to nine months for human transcription of the 1940 Census.<sup>70</sup> The Material Culture of Wills project also uses HTR to rapidly transcribe 25,000 wills from the Prerogative Court of Canterbury.<sup>71</sup> This model will later enhance the National Archives (UK) manuscript collection's accessibility, providing genealogists with digital access to more resources.<sup>72</sup>

AI innovations in primary source management have transformed the efficiency of data retrieval and organisation in archives and opened new datasets for genealogical and historical research.<sup>73</sup> Gibbs notes that digitised records enable new research techniques and dissemination methods, enhancing data analysis and interpretation.<sup>74</sup> However, Crymble recommends the establishment of digital archives with multidisciplinary teams to integrate AI technology while maintaining the integrity of the data.<sup>75</sup>

This transition to digital archives requires new management and research approaches to interpret the growing digital data volume.<sup>76</sup> Johnson and Bell advocate modernising current data classification systems, which are based on old card cataloguing systems. (Fig.6)



*Figure 6 The National Library of Australia's shift from a card catalogue to a computer system.<sup>77</sup>*

Bell suggests adopting an associative approach using Natural Language Processing (NLP) for improved retrieval.<sup>78 79</sup> NLP facilitates searches using everyday language, providing more expansive and contextually relevant results.<sup>80</sup> However, my research findings indicated that respondents were only moderately impressed with NLP search capabilities in their current format and preferred to continue using traditional methods.

Sherratt and Crymble assert that the evolution from microfiche to digital imaging and AI indexing has democratised access to historical materials, offering researchers unparalleled ease and efficiency.<sup>81 82</sup> They emphasise that AI's potential to refine data organisation and analysis will unlock more profound insights into historical narratives.<sup>83</sup> The impact of technology on historical research, as described by Crymble and Sherratt, is showcased in the three projects investigated during my research.<sup>84 85</sup>

Building on these insights, Milligan argues that leveraging AI requires training researchers and genealogists to understand digital research principles and the use of AI technology.<sup>86</sup> He believes understanding AI's impact on research choices is crucial for effective data collection, analysis, and historical interpretation.<sup>87</sup>

Godfrey and Cowley also highlight the importance of robust contextual databases in historical research.<sup>88 89 90</sup> AI tools and digital methodologies can enhance the depth and breadth of historical analysis and understanding of this research's projects; Godfrey's work used AI to link historical records, analysing the impact of solitary confinement on convict life

expectancy.<sup>91</sup> His AI-enabled spatial analysis enriched historical criminology by connecting the data to locations, offering insights into penal transportation.<sup>92 93 94</sup>

However, the use of AI in historical research raises significant ethical considerations, with Johnson, Gibbs, and others acknowledging the inherent biases in digitisation. Sherratt's statement, 'collections are formed by exclusion, by decisions about whose lives and voices matter', underscores the potential for choices about what to digitise, indexing methods, and access protocols to subtly influence historical narratives. Little further highlights these ethical dilemmas, particularly when digitised data is used to train AI systems. These authors stress the critical importance of strong ethical frameworks, human validation and ensuring transparency in all research projects.

These technological advancements are fundamentally changing genealogical research. The integration of AI necessitates re-evaluating traditional approaches and adopting new strategies to manage and interpret digital data. Ongoing research is needed to fully comprehend AI's impact. This study aims to contribute insights into AI's impact on genealogy.



## 2.3 Adaption and Change Management in Genealogy

Over time, genealogists have adapted to technological changes, from microfiche to digital databases and online genealogy platforms, demonstrating a willingness to integrate new technologies, including AI, into traditional practices. (Fig.7)

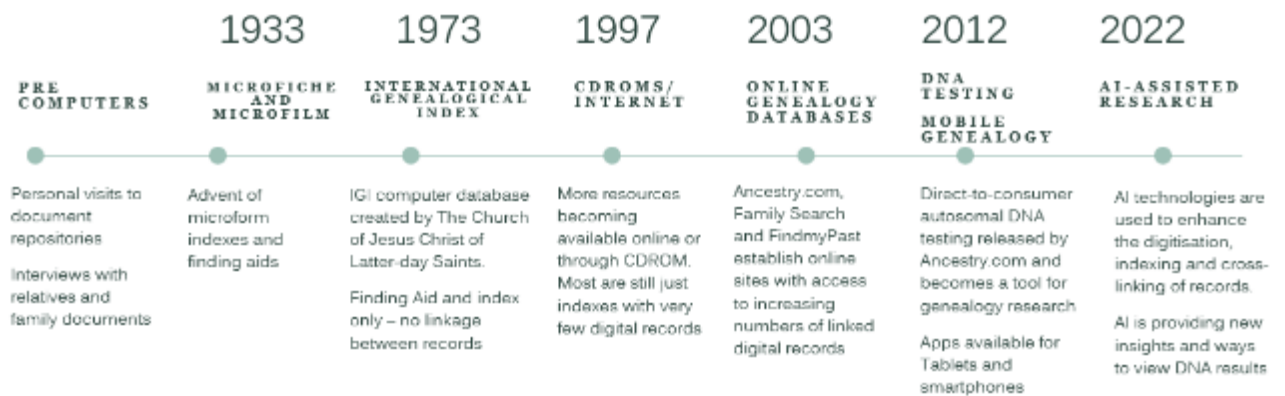


Figure 7 Genealogical research methods timeline.<sup>95 96 97</sup>

Rogers' Diffusion of Innovations theory provides a framework for understanding technology adoption, identifying five adopter categories: Innovators, Early Adopters, Early Majority, Late Majority, and Laggards.<sup>98</sup> (Fig.8) This segmentation aids in understanding how societies embrace new technologies and develop effective change management strategies.<sup>99</sup> Most of my testing cohort are Innovators/ Early Adopters, typically more open to new technologies, and their potential positive bias must be considered during analysis.<sup>100</sup>

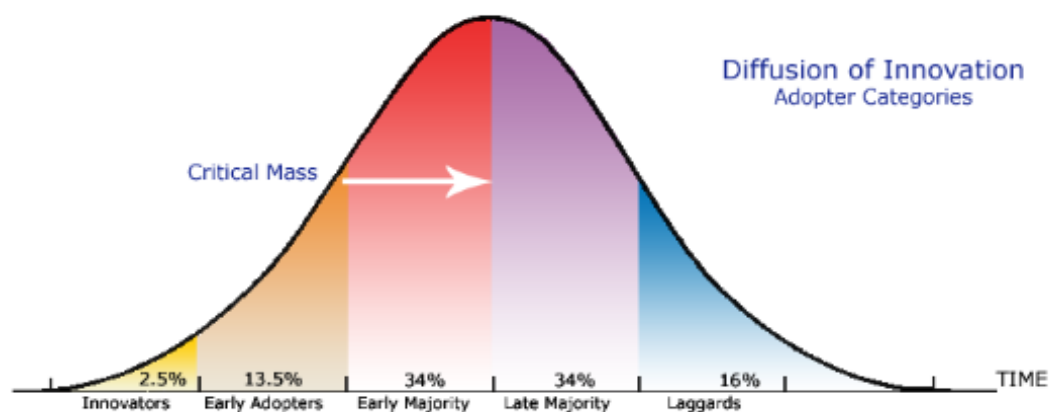


Figure 8 Rogers' Theory of Diffusion of Innovation.<sup>101</sup>

Understanding the behaviours and needs of genealogists can also inform how new technologies are integrated into research practices. Duff and Johnson's research explores genealogist's information-seeking behaviours, detailing research stages, tools, expertise, and challenges.<sup>102</sup> Their findings emphasise the need for tailored archive systems, a need that has grown with AI's integration. Their 2002 and 2003 studies demonstrated genealogists' adaptability to early internet resources despite scarce online primary sources, showing their resourcefulness and relationship with technology.<sup>103 104</sup>

Building on earlier works, Lucy investigated technological integration within genealogy research.<sup>105 106 107 108</sup> Her study revealed a reliance on technology and the internet for research, moving away from traditional, physical archive-based methods.<sup>109</sup> She highlighted genealogists' proficiency in using various resource formats, concluding they wanted faster digital searches and were open to technological changes.<sup>110</sup> Her methodology informed the development of data collection methods used in this study.

The genealogical community is only beginning to understand generative-AI's impact. Bettinger highlights its potential benefits and limitations, suggesting generative-AI useful for content generation and organisation, but cautions against using it for complex analytical reasoning or factual information.<sup>111</sup> He advocates for a measured approach, urging genealogists to use generative-AI judiciously while staying informed about its potential and limitations.<sup>112</sup> The lack of information on generative-AI's use in genealogy has seen the community use social media for real-time guidance and best practices.<sup>113</sup> AI genealogy pioneers, like Bettinger, Little, Hoitink and Leeds, use Facebook and blogs to explore generative-AI, sharing breakthroughs and challenges.<sup>114 115 116 117</sup>

Genealogists have shown a remarkable ability to adapt to change and embrace new technologies, strongly suggesting they are open to AI's potential. This research aims to provide empirical evidence on generative-AI's capabilities in genealogy and the community's willingness to adapt.

## 2.4 Role of AI in Genetic Genealogy

DNA analysis has revolutionised genealogy, reshaping our understanding of family connections. This transformation is driven by AI's ability to process complex genetic data swiftly and accurately. This section explores AI's role in genetic genealogy and addresses some associated ethical implications, focusing on AI's contributions rather than genetic genealogy methods and practices.

Caine claims AI's proficiency lies in handling the vast, intricate datasets typical of genetic research, bringing precision to interpreting genetic data, and aiding in significant breakthroughs.<sup>118</sup> AI's contribution ranges from analysing SNP microarrays to facilitating advanced genetic testing and enhancing the understanding of ancestry and family connections.<sup>119</sup> (Fig.9)

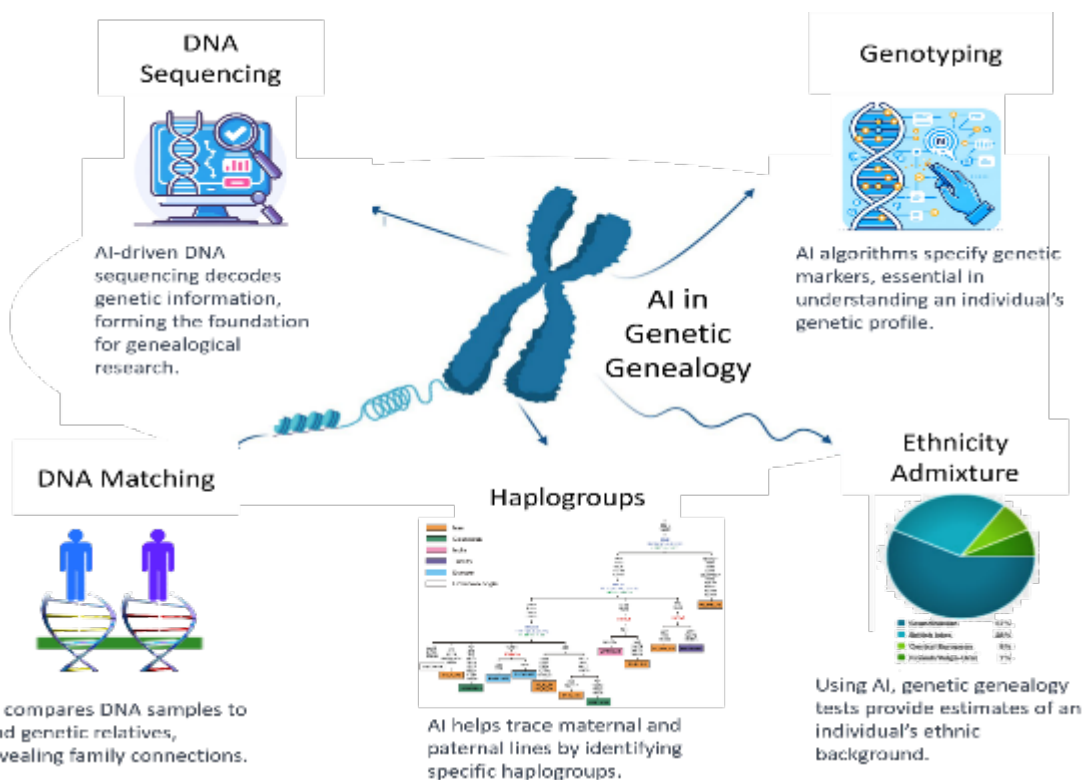


Figure 9 Use of AI technologies in genetic genealogy.<sup>120</sup>

## **DNA Testing**

Glynn's 2022 report highlights that over 41 million people worldwide had taken direct-to-consumer DNA tests, using AI to analyse DNA from samples using high-density Single Nucleotide Polymorphism (SNP) microarrays, pinpointing them accurately.<sup>121 122</sup> Dias explains AI neural networks compare test data against reference panels using technologies like computer vision, mapping algorithms, and sequence analysis to detect genetic variants with high precision.<sup>123 124</sup>

Hanacek notes each testing company's SNP chips and proprietary AI algorithms analyse public databases, reference panels, and their customer results database, influencing the outcomes and causing variances in interpretation and accuracy between companies.<sup>125</sup>

## **DNA Matching and AI**

Holton explains genotyping provides ethnic admixture estimates and precise haplogroup details, essential for tracing ancestral lineages.<sup>126</sup> Alharbi describes the role of machine learning and neural networks in this process, which performs intricate calculations to extract meaningful patterns from large genomic datasets.<sup>127 128</sup> Holton emphasises AI-enhanced data's role in accurately determining haplogroups, especially with the influx of data from Next-Generation Sequencing tests, providing new insights into mtDNA and Y-DNA lineages.<sup>129</sup>

Holton and Stankus caution against overreliance on admixture results, which depend on the diversity and size of reference groups for accurate estimates.<sup>130 131</sup> For example, AncestryDNA uses a reference panel of about 71,306 samples from 88 regions, continually refining its machine-learning algorithms as its database expands to improve precision.<sup>132</sup>

AI in DNA analysis, especially in identifying cousin matches, is an asset for genealogists. AI's machine-learning algorithms analyse DNA segments and identify genetic relatives from extensive databases and are one of the most effective tools for expanding genealogical research.<sup>133 134</sup> Hu and Ball discuss the sophisticated algorithms and estimators behind many features offered by testing companies, illustrating AI's capacity to empower genealogists.<sup>135 136 137</sup> Bettinger notes AI-enabled DNA matching facilitates discovering unknown family branches and deepens understanding of genetic linkages, creating more precise family histories.<sup>138</sup>

## **Impact of Genetic Genealogy**

Stallard discusses AI-enabled genetic genealogy's impact on traditional research methods, revealing the objective nature of DNA data, its importance as a verification source and the potential to share data across platforms, means DNA is a unique resource.<sup>139</sup> Holton's work in the Battle of Bannockburn and the Treaty of Arbroath projects demonstrates how AI-enabled genetic testing combined with traditional sources can reveal new insights into medieval ancestries.<sup>140</sup>

## **Challenges and Ethical Considerations**

Christofides raises ethical concerns about privacy and data security in genetics, warning of the risks of unauthorised access, misuse of genetic data and undisclosed use or sale of raw data for research.<sup>141</sup> Similarly, Parikh highlights biases in AI from the quality of training data or disproportionate ethnic representation in reference panels, distorting genetic analysis.<sup>142</sup> Ménard emphasises the importance of rigorous validation of AI's accuracy and interpretation, addressing concerns about AI's susceptibility to errors and biases.<sup>143</sup> Their call for establishing scientifically robust, ethical and socially sensitive practices to guide testing aligns with Royal.<sup>144</sup>

The Genetic Genealogy Standards and ISOGG's ethical framework address these ethical challenges, ensuring genetic genealogy adheres to principles of privacy, accuracy, and equity, balancing technological advancements with moral responsibility.<sup>145 146</sup>

Given proprietary AI's prevalence in genetic genealogy, outside genealogists' control, my study concentrated on understanding its influence on research practices and exploring potential applications of end-user AI tools.

## 2.5 With Great Power comes Great Responsibility – Ethically using AI in Genealogy

Mollick, referencing Kratzenberg's Law, emphasises AI's impact is neither inherently positive, negative, nor impartial; but depends on its application.<sup>147 148</sup> He claims AI's introduction heralds significant workplace shifts that cannot be ignored.<sup>149</sup> This perspective is relevant to genealogists and reflects the importance of deliberately integrating ethical considerations into AI-assisted genealogical research, ensuring that technology enhances rather than detracts from genealogical integrity.

The unprecedented speed and scope of the change arising from the release of generative-AI presents unique challenges and opportunities. Governments and organisations worldwide are grappling with integrating ethical considerations into the AI paradigm.<sup>150</sup> Fjeld mapped global ethical frameworks, identifying eight key themes: privacy, accountability, safety and security, transparency and explainability, fairness and non-discrimination, human technology control, professional responsibility, and promotion of human values.<sup>151</sup> These themes are echoed in the IEEE's guidelines and the EU's AI Act.<sup>152 153</sup>

AI developers also play an essential role in the ethical use of AI by embedding ethical standards into algorithms and providing tools for bias mitigation. Complementing these technical efforts, Li and Mollick advocate for a human-centred approach, emphasising AI systems that uphold human dignity, safety, and inclusivity, which reflect core genealogical values.<sup>154 155</sup>

Mollick emphasises the practical application of these principles in organisational settings, arguing that fostering an understanding of ethical AI across various sectors is necessary for its effective implementation.<sup>156</sup> Adopting AI has raised similar challenges in genealogy, which are increasingly important to address.

MacEntee lists genealogists' ethical AI concerns as intellectual ownership, privacy, transparency, societal biases, and misinformation.<sup>157</sup> Cowgill expands, stressing the importance of high-quality input data to mitigate biases, especially gender and race biases and the necessity of user education to reduce the risk.<sup>158</sup>

Acknowledging generative-AI's limitations in factual accuracy, which is critical in genealogy, Bettinger suggests carefully considering its use for factual research.<sup>159</sup> He also addresses the legal and ethical complexities of AI-generated content, questioning if AI can commit plagiarism, as, to date, only humans are recognised in this context.<sup>160</sup> This issue is particularly relevant in genealogy, where the authenticity and originality of genealogical narratives are integral. Bettinger calls for transparency in the use of AI tools, suggesting that genealogists should disclose AI assistance, aligning with publication guidelines for academic journals.<sup>161 162 163</sup>

Little's balanced approach to ethics, disclosure, and privacy aligns closely with these arguments.<sup>164</sup> He advocates for pragmatic AI use, suggesting AI assistance be regulated similarly to human assistance, and cautions against sharing sensitive data with generative-AI.<sup>165</sup> Like Mollick, MacEntee and Bettinger, his guiding principles encourage a human centric approach to AI, highlighting the importance of addressing assumptions and engaging interactively with AI technologies.<sup>166 167 168 169</sup>

These issues reflect broader global challenges in ethical AI development and the need to determine the fundamental principles for responsible AI use in genealogy. Professional genealogy organisations promote ethical frameworks that emphasise thorough, verifiable research, accuracy, and ethical information handling that align closely with these principles.<sup>170 171 172 173</sup> These frameworks are being adapted to address AI's unique challenges within genealogy, with grassroots initiatives, like the AI and Genealogists Facebook group's decision to watermark AI-generated images, showcasing a community commitment to ethical AI use.<sup>174 175</sup>

This blend of established frameworks and genealogist-led innovations creates a comprehensive approach to integrating AI into genealogical research. This ensures practices remain consistent with genealogy's core values of accuracy, transparency, and integrity.

One of my research aims is to fill the knowledge gap on the ethical use of AI-assisted research for genealogy by focusing on how genealogists can incorporate ethical guidelines into their AI-assisted research methodologies.

## **2.6 Literature Review Conclusion**

Based on the literature review insights, a theoretical framework for understanding AI's role in genealogy can be developed, guided by three components: generative-AI's capabilities in genealogy workflows, ethical integration of AI and the broader impact of AI integration on genealogy research. The following chapter uses this framework to outline the methodologies used to gather empirical evidence, detailing the research strategy, data collection techniques and sample selection criteria.



## Chapter 3 – Research Methods

This chapter builds on the theoretical framework established in the Literature Review to gather empirical data on AI's role in genealogy workflows. The diverse research questions required a multi-method approach to capture the AI's multifaceted impact. This chapter outlines specific methodologies for each research question, covering data collection techniques, ethical considerations, and analytical strategies and concludes with methods used to mitigate methodological challenges and limitations.

The theoretical framework aligns with the six research questions, highlighting the research's significance within the broader academic discussion and providing a foundation for exploring AI's application in genealogy. (Table.1) Integrating the research questions within the framework allowed a detailed examination of AI integration in genealogy.

Theoretical Framework	Research Questions
Generative-AI's capabilities in genealogy workflows.	<ol style="list-style-type: none"> <li>1. What are the benefits and limitations of generative-AI for genealogy purposes?</li> <li>2. How can generative-AI be used in genealogy, how effective, efficient and accurate is it for genealogical purposes and how aware are genealogists of these tools?</li> </ol>
Ethical integration of AI into genealogy workflows.	<ol style="list-style-type: none"> <li>3. What ethical considerations and safeguards are needed when using AI and what risks arise from not implementing them?</li> </ol>
The broader impact of AI integration on genealogical research.	<ol style="list-style-type: none"> <li>4. How is AI integrated into current genetic genealogy processes, and what are the benefits and challenges?</li> <li>5. How is AI changing the availability, access and use of historical documents?</li> <li>6. How is AI being used to re-interpret historical data to provide new insights for genealogical research?</li> </ol>

Table 1 Correlation between the theoretical framework and the research questions.

### 3.1 AI's Capabilities in Genealogy (Research Questions 1-5)

#### Research Strategy

A mixed-method approach was adopted to investigate Research Questions 1-5.<sup>176</sup> This strategy enabled a comprehensive assessment of AI's efficiency, effectiveness, and accuracy by integrating qualitative and quantitative data, aligning with Creswell's recommendation for studying complex problems by combining statistical information with the voices of individuals.<sup>177</sup>

Inspired by Hörnemalm and Dell'Acqua's research into ChatGPT in the workplace, a series of Capability Analyses (CA) were developed.<sup>178 179</sup> These benchmark-style analyses were designed to quantitatively assess AI's efficiency and accuracy while gathering qualitative feedback from genealogists on AI's effectiveness. This convergent approach comprehensively explored AI's benefits and limitations in genealogical research.<sup>180</sup>

This methodology evaluated AI's impact on genealogical research through experimentation within participants' real-world work environments for realism and relevance. The approach aligned with Creswell's recommendation that experimentation provided quantitative rigour and qualitative depth and Mollick's observation that 10 hours of experimentation was the best way to master generative-AI.<sup>181 182</sup> Additionally, following Rogers' assertion that early adopters serve as educators and advocates, the research strategy encouraged participants to embrace this role.<sup>183</sup>

#### Data Collection - Design Phase

Ethical Approval: Approval was granted by the CLL Ethics Committee, as living people were involved in this research.

The design phase began with a focus group of experienced genealogists from FFACT's Education team, who identified genealogical tasks for the CAs to ensure real-world relevance. Using an independent group prevented bias from my experience, enabling objective task selection for evaluating generative-AI's capabilities.

The tasks chosen represented a wide range of genealogical activities, from simple narratives and extraction of information to more complex challenges in analysis and data interpretation. Many tasks fell into the categories identified by Little and Mollick, where generative AI performs well, such as summarisation, extraction, generation, and translation.<sup>184 185</sup> Other

tasks, such as creating a family tree, fell into areas Dell’Acqua described as the “jagged frontier”. These are areas where generative AI would likely be challenged to produce a meaningful response.<sup>186</sup> Table.2 provides an overview of the developed CAs; more details are available in Appendix.A

Capability Analysis	Area of focus	Goal
CA1	Brainstorming and Narrative Generation	Evaluation of generative-AI’s capability to generate topics and content for websites/blogs.
CA2	Extraction	Evaluation of the effectiveness of generative-AI in extracting relevant information from diverse genealogical records.
CA3	Translation	Evaluation of generative-AI’s accuracy in translating and contextualising non-English documents.
CA4	Data Interpretation	Evaluation of generative-AI’s ability to identify relationships and inconsistencies and develop hypotheses from genealogical data.
CA5	Analysis and Data Extraction of complex data	Evaluation of generative-AI’s skill in extracting and analysing information from complex documents such as a will and creating a family tree from the information.
CA6	Ethics	Investigation of AI’s capacity to handle ethical concerns in genealogy and the ethical use of AI for genealogists.
CA7	Genetic Genealogy	Assessment of the impact of integrating genetic genealogy into conventional research strategies and the use of interpretive AI tools.
CA8	Document Research and Planning	Assessment of AI’s ability to assist in locating documents and planning research activities

*Table 2 Structure and content of the Capability Analyses.*

The CAs were designed to evaluate generative-AI’s ability to perform specific tasks rather than assess the genealogist’s performance. As illustrated in Fig.10, the CAs collected data on the response from generative-AI and did not collect information on work created by the respondents.

# Capability Analysis Test 4: Data Interpretation

Objective: To assess generative AI's ability to accurately analyse genealogical data to identify relationships, inconsistencies, and hypotheses. This is a common but complex genealogical task requiring skills in data matching, data verification and experience in understanding common data anomalies.

## Your Tasks

1. **Error Detection (You and AI):** Review the dataset for any inconsistencies or errors in dates, names, or relationships. Ask AI to explain any discrepancies in ages or birth/death years across different documents in the dataset.
  2. **Relationship Analysis (AI):** Determine the specific family connection of Elizabeth Johnson, listed in the 1930 census, to the SMITH family.
  3. **Theory Formation (AI):** Formulate a theory about the unknown father of Sarah Johnson SMITH based on the available information and suggest the next step to test this theory.
- NOTE:** This task is a complex one and is currently right on the edge of LLM's capabilities. This will change as LLMs continue to develop. To get the best response, you should act as the human moderator and check and challenge the responses it provides.
- Help:** A set of prompt suggestions for the task is available at the end of the document.

## Responses - Data Interpretation

Researcher ID: \_\_\_\_\_

Task 1: Error Detection (You and AI)					
1.1 Review the dataset for any inconsistencies or errors in dates, names, or relationships, then ask the AI app to do the same task. Ask AI to explain any discrepancies in ages or birth/death years across different documents in the dataset. Challenge any AI responses you don't agree with and collaborate to produce a final list.					
Which AI app did you use?					
Evaluate the result using the evaluation criteria. Score 1-5 1=(Lowest) - 5=(Highest)					
	1	2	3	4	5
How accurately did the AI pick up data anomalies?					
Did it pick up any significant ones that you missed?					
Did it miss any significant anomalies that you identified?					
Was it able to provide explanations for the anomalies it picked up?					
How well did the system respond to your input and feedback?					
Comments: Provide your feedback and comments on the error detection process. Note if you used any aids such as a list or spreadsheet.					

Figure 10 Example of Capability Analyses.

This approach helped ensure the study centred on the effectiveness of generative-AI, and eliminated any perception the respondents were being tested. The participants only used their work to verify the accuracy and relevance of generative-AI's output, ensuring an unbiased evaluation of its capabilities.

Each CA provided information on:

- the CA's objective,
- the parameters,
- the evaluation criteria,
- a list of tasks,
- test data where needed,
- the response form for the quantitative questions (Likert scales and time) and qualitative questions (Open-ended feedback), and
- sample prompts.

### **Data Collection Phase:**

Participants were drawn from FHACT, the only genealogical society in the Australian Capital Territory (ACT), with around 600 members. My association with FHACT, its 2023 workshops on generative-AI, and culture of community support and ongoing education, made it an ideal choice for this research.

A self-selected sample of 21 Society members expressed interest in the study, with 16 participating in the project briefing session and AI workshop. Non-attendance was attributed to work, travel, and health commitments. However, one prospective participant withdrew over concerns that AI could perpetuate sexist and racist stereotypes, a relevant consideration in the wake of public debates around responsible AI use.

Although more participants might yield more statistically robust results, expecting large numbers is unrealistic at this stage of generative-AI technology, which is still in the early adoption phase based on Rogers' Diffusion of Innovation theory.<sup>187</sup> This phase represents 2.5-5% of the population, equating to 15-30 participants for FHACT. The study's sample of 16 participants, 2.1% of FHACT's membership, is a realistic sample size for this phase of technology adoption. The small size of the test group and my connection to FHACT facilitated participant comfort and made the sample group easy to access and manageable.

An introductory AI workshop inspired by Dell'Acqua's findings that proficiency enhances task performance prepared participants.<sup>188</sup> This step was crucial for familiarising them with AI applications, and improved the reliability of findings. Participants received the CAs and had one month to evaluate AI's performance in their typical working environment. They could choose the generative-AI model they used for the evaluation. This approach allowed data collection in a low-stress, realistic setting with their own equipment.

A preliminary survey collected demographic information and initial AI perceptions of AI, contextualising the analysis. A debriefing survey followed the completion of the CAs to measure changes in respondent's perceptions of AI's effectiveness.

## Demographics of Testing Cohort

The demographic profile of the testing cohort is important as it provides a backdrop against which the data on AI-integrated genealogy workflows can be evaluated and bears implications for the research findings.

The testing cohort was representative of many genealogists who, although not paid to research, approach their work with a professional mindset. While all genealogists can benefit from generative-AI, this group has more autonomy in choosing whether or not to use it, unlike professional and academic genealogists whose AI use is often dictated by workplace policies. This autonomy enables an unbiased evaluation of generative-AI's capabilities, reflecting genuine choices and experiences of genealogists who can independently decide to integrate AI into their workflows.

The preliminary survey revealed that most participants were experienced genealogists, with the majority rating their skill level as intermediate or above. An intermediate level was defined as having a solid grasp of genealogical research methods. The cohort included one beginner and one professional genealogist. The value of having experienced genealogists is significant, as their expertise allows them to better evaluate the accuracy and relevance of generative-AI responses, leading to more reliable assessments of the technology's capabilities. Most participants rated their computer skills as intermediate or higher, minimising the impact of poor computer skills on the study's results, ensuring the focus remained on evaluating the effectiveness of generative-AI. (Fig.11)

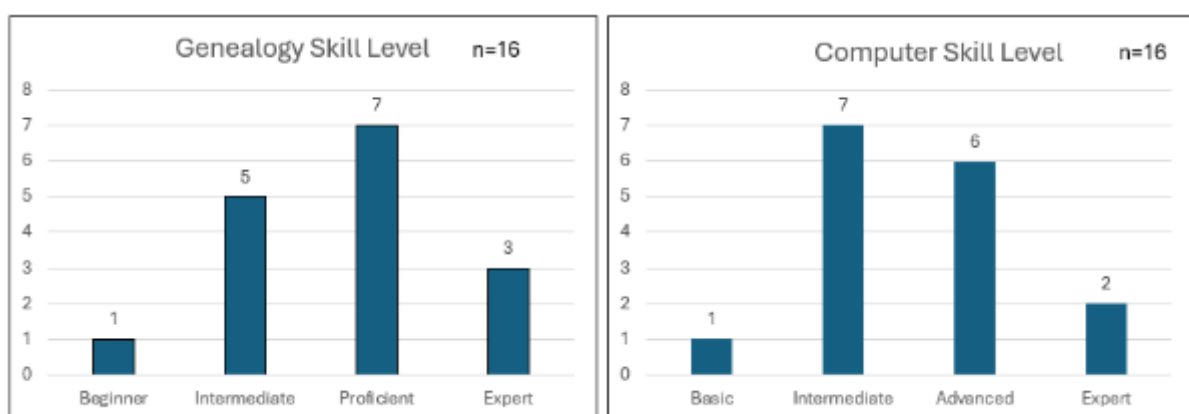


Figure 11 Genealogical and computer skill level of the testing cohort.

The preliminary survey also revealed that the cohort was between 60 and 80 years old, which is representative of many Australian genealogical societies with aging memberships. (Fig.12)

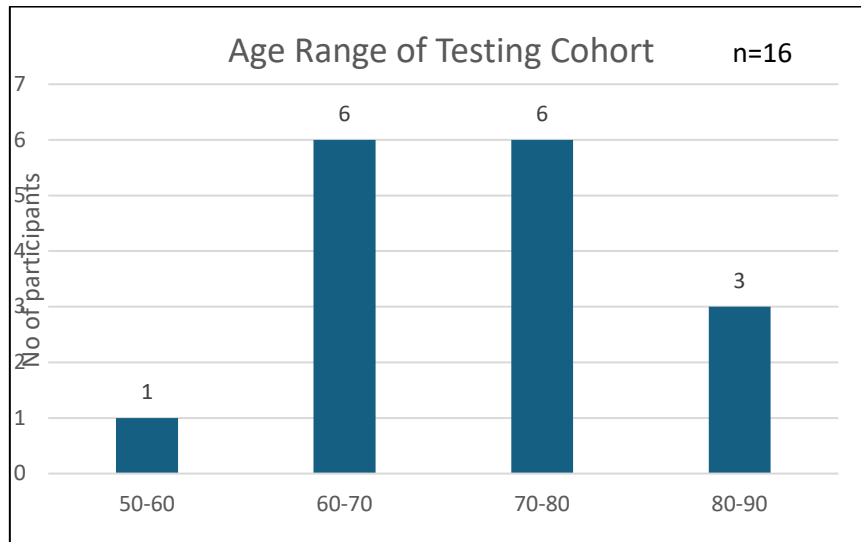


Figure 12 Age of testing cohort.

There was a notable enthusiasm for AI integration into genealogical research, with 87% of participants expressing enthusiasm, reflecting a positive bias towards technological innovation. This enthusiasm is significant given the cohort's age demographic, challenging the stereotype of digitally challenged older adults and strongly suggests they fall into Rogers' innovator/early adapter group.<sup>189</sup>

## Data Analysis

Collected data was analysed using an integrated approach.<sup>190</sup> Quantitative data, including performance metrics and task completion times, was statistically analysed using Excel and OpenAI's Data Analyst. Tools such as Survey Monkey facilitated the organisation and analysis of survey responses, and spreadsheets were used to process and visualise the quantitative data. Concurrently, qualitative feedback was thematically examined to provide insights into generative-AI's perceived effectiveness and limitations. Using Creswell's methodology, an inductive coding system identified themes and patterns in the data without preconceived categories.<sup>191</sup> Data was systematically reviewed, with codes developed based on the content, then refined and grouped into broader themes to derive meaningful insights. (Fig.13) By merging the qualitative and quantitative results, a comprehensive understanding of AI's role in genealogy was achieved, highlighting areas of strength and opportunities for improvement.<sup>192</sup>



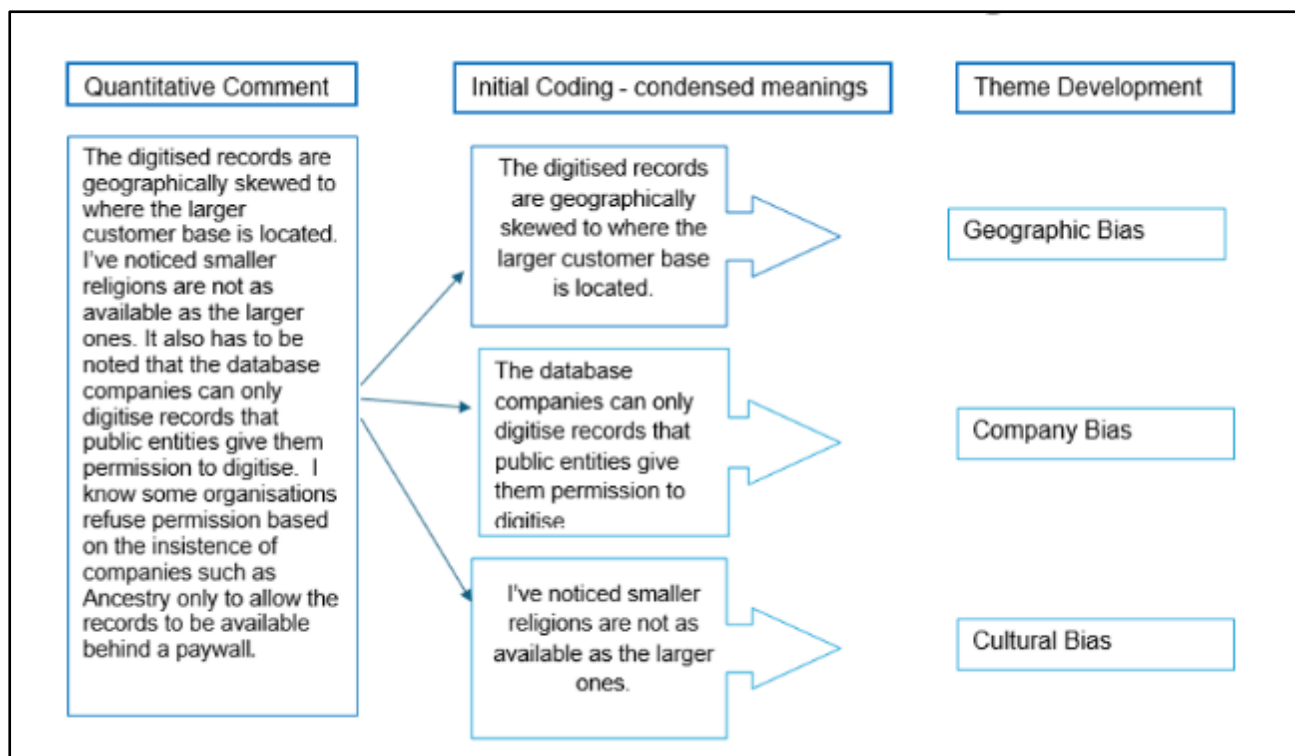


Figure 13 Example of steps for inductive coding of quantitative data.<sup>193</sup>

## Additional Methodological Considerations for Research Questions 3-5

### Ethical Integration of AI into Genealogy Workflows

**RQ.3 What ethical considerations and safeguards should be implemented when using AI, and what are the risks associated with not implementing them?**

This question explores the ethical implications of AI in genealogy research through an analytical review of existing literature and qualitative and quantitative data. CA6 was designed to critically assess AI tools' capacity to navigate and address ethical dilemmas inherent in genealogical research. It also examined how AI applications could assist genealogists in ethical decision-making processes and how they align with established ethical frameworks identified during the literature review phase.

Preliminary and debriefing surveys captured changes in participants' ethical concerns regarding AI use.

## **The Broader Impact of AI Integration on Genealogy Workflows**

***RQ.4. How is AI integrated into current genetic genealogy processes, and what are the general benefits and challenges observed?***

***RQ.5. How is AI changing the availability, access and use of historical documents?***

Given AI's pivotal role in these fields and the increasing reliance on AI technologies for research, CA7 and CA8 were designed to collect information on the broader integration of AI into genealogists' work practices. CA7 investigated the effects of integrating genetic genealogy into conventional research strategies and the use of interpretive AI tools provided by genealogy companies. CA8 focused on genealogists' search strategies and responses to Natural Language searching, aiming to test specific concepts related to genealogists' practices, as raised by Lucy.<sup>194</sup>

## 3.2 The Impact of AI Integration on Genealogists: RQ 6

### ***RQ.6. How is AI being used to re-interpret historical data to provide new insights for genealogical research?***

A different research approach was adopted to evaluate AI's use to uncover new insights from historical datasets and contribute new information on our ancestors' lives.

#### **Research Strategy**

Using Yin's methodological framework, a qualitative approach examined three real-world projects re-interpreting Australian convict records with AI technologies.<sup>195</sup> This approach was selected for its strength in enabling in-depth exploration of complex phenomena within their real-life contexts.<sup>196</sup> The projects covered a range of AI technologies, addressing the research question through varied examples of technological application in historical research.

#### **Data Collection**

Following Denscombe's advocacy for deep narrative understanding, data was collected through interviews with project leaders, supplemented by academic papers and presentations.<sup>197</sup> Semi-structured interviews allowed for targeted questions related to research objectives, such as inquiries about handling data bias, data quality requirements, AI's role in data reinterpretation and open-ended discussions to capture a broader understanding of each project's challenges and achievements.<sup>198</sup> This approach ensured a rich data collection, combining experience, theoretical knowledge and practical insights into A application in historical and genealogical research.

The projects were selected because of the significance of convict research in Australian genealogy, the involvement of genealogists in the projects, and the close relationship between the project leaders and the Australian genealogical community. The projects are:

- *Ironclad Sisterhood*: A re-examination of an overlooked dataset on female convicts held by a genealogical society in New South Wales.
- *Landscapes of Production and Punishment*: The re-interpretation of the movements and activities of Tasmanian convicts using LiDAR and AI-enabled spatial analysis.
- *Making Crime Pay*: The use of AI to manage large datasets of Tasmanian convict records, following their experiences from conviction to death. The project also developed an AI-enhanced immersive exhibition, *Unshackled*, to share the convicts' stories with the broader community.

## Data Analysis

Data analysis proceeded project-by-project, using an inductive coding system to identify common themes across interviews. This thematic analysis was enhanced by integrating insights from academic and presentation materials, providing a detailed understanding of each project's contribution to genealogical research. Key themes included data bias, the ethical use of AI, and the innovative presentation of historical information.

### 3.3 Limitations and Mitigation Strategies

While the main research strategies chosen, mixed methods and interviews, are comprehensive and robust, like all methodologies, there are inherent limitations and potential hurdles.

**Sample Representation and Generalisability of Findings:** The small testing cohort does not represent the broader genealogy population, limiting the finding's generalisability. This also applies to project interviews, as the projects may not describe all AI-enabled projects. Mitigation: Future studies should widen the participant pool to include younger researchers, novices, and professional genealogists for better representativeness and generalisability. Broadening the investigation into other AI-enabled projects is also recommended.

**Bias in Participant Selection:** As the participants self-selected for the study, there is a risk of selection bias. Those who are more technologically savvy or have a pre-existing interest in AI are likelier to participate, potentially skewing the results.

**Participant Bias:** Rogers suggests participants likely fall into the innovator/early adopters categories, which are more receptive to new technologies.<sup>199</sup> This predisposition was reflected in the participants' overwhelmingly positive response to AI in the preliminary survey, potentially skewing the findings to a more favourable perception of AI's role in genealogy.

Minimising Biases: The initial biases tied to the early phases of technology adoption and the associated technological enthusiasm were challenging to mitigate in the small testing pool. As suggested by Creswell, these biases were acknowledged, and over-generalisation was carefully avoided in the findings.<sup>200</sup> Future studies, as the use of AI in genealogy matures, may offer different insights as perceptions and adoption rates change.

**Rapid Evolution of AI Technologies:** AI is rapidly evolving, with new tools and applications emerging continuously. The CAs capture the state of these tools at a specific point in time and do not reflect future advancements or limitations, potentially dating the findings.

Mitigation: Continual Review: Although the CAs are a snapshot of the technologies available in February 2024, an ongoing review of the literature and innovations throughout the research ensured that later AI developments were integrated into the dissertation wherever possible, ensuring the research remained current.

By acknowledging and mitigating these limitations, this study provided accurate, representative, and timely insights into AI's use in genealogy. Future research will benefit from this learning, potentially offering a deeper understanding of the field as it evolves.

### **3.4 Methodology Conclusion**

This chapter has outlined the research's methodological approach, aligning the mixed methods strategy with the research questions. Qualitative and quantitative analyses thoroughly investigated AI used in genealogy to empower genealogists. The empirical data will help fill knowledge gaps, address real-world challenges and contribute to a meaningful understanding of AI's role in this evolving field. The analysis of the results will be discussed in the next chapter.

## Chapter 4 – Empirical Findings: Description, Analysis & Synthesis

This chapter uses the theoretical framework to explore the findings for each research question. (Table.3) Each section begins with a data overview, followed by empirical findings and a comparison with the literature review insights. Within each question, the ongoing analysis builds on the results, enhancing the understanding of the broader implications of the data. Chapter 5 will provide a comprehensive conclusion, with a summative reflection on all research questions and outlining the study’s insights.

Chapter Section	Theoretical Framework	Research Questions
4.1	Generative-AI's capabilities in genealogy workflows.	<ol style="list-style-type: none"> <li>1. What are the benefits and limitations of generative-AI for genealogy purposes?</li> <li>2. How can generative-AI be used in genealogy, how effective, efficient and accurate is it for genealogical purposes and how aware are genealogists of these tools?</li> </ol>
4.2	Ethical integration of AI into genealogy workflows.	<ol style="list-style-type: none"> <li>3. What ethical considerations and safeguards are needed when using AI and what risks arise from not implementing them?</li> </ol>
4.3	The broader impact of AI integration on genealogical research.	<ol style="list-style-type: none"> <li>4. How is AI integrated into current genetic genealogy processes, and what are the benefits and challenges?</li> <li>5. How is AI changing the availability, access and use of historical documents?</li> <li>6. How is AI being used to re-interpret historical data to provide new insights for genealogical research?</li> </ol>

Table 3 Mapping of research questions to the theoretical framework.

### 4.1 Generative-AI’s Capabilities in Genealogy

**Research Question 1 (RQ1): What are the benefits and limitations of generative-AI for genealogy purposes?**

**Research Question 2 (RQ2): How can generative-AI be used in genealogy, how effective and efficient is it, and how aware are genealogists of these tools?**

Generative AI, particularly LLMs, is a powerful yet user-friendly tool accessible to anyone proficient in reading and writing.<sup>201</sup> To understand its capability for genealogical activities, CA1-8 collected qualitative and quantitative data. This section analyses RQ1 and RQ2 concurrently, beginning with the uses of generative-AI, then examining its benefits and limitations, and providing insights into its effectiveness for genealogy research.

## Framework and Scope of Assessment

1. This study focused on free generative-AI capabilities to minimise any cost to the testing cohort. Due to subscription requirements, text-to-image tools like DALL-E.3 or Midjourney and Transkribus, a handwriting analysis tool, were not tested as part of the CAs.
2. The respondents used various generative-AI models for CA evaluation. ChatGPT-3.5 was the most popular among the free options due to its performance and higher security levels.<sup>202</sup> Some chose ChatGPT-4 for its advanced features, better performance and lower hallucination rate.<sup>203</sup> Others switched between platforms to achieve optimal results. This variety in model usage highlights the individual needs and preferences of genealogists. (Table.4)

Generative AI Model	No. of Respondents using Model
ChatGPT-3.5	14
ChatGPT-4 (subscription)	2
Perplexity	1
Microsoft Copilot	2
Bing	2
N.B. Claude was not available in Australia in Feb 2024	

*Table 4 Use of Generative-AI platforms for CA evaluation.*

3. Since data collection, generative-AI capabilities have evolved. For example, in February 2024, only ChatGPT-4 allowed document uploads for analysis. By May 2024, these features were extended to free users with the release of ChatGPT-4o, marking a significant improvement in accessibility.<sup>204</sup>
4. Generative-AI is not a search engine nor a substitute for traditional databases, although some models can access the internet. Its primary function remains generating content based on learned patterns from the training data, not retrieving information from the internet. While it can suggest theories or narratives based on existing data, it cannot directly answer specific queries with new information. For example, searching for a birth record may result in plausible but fictitious information.

## Uses of Generative-AI in Genealogy

Genealogy, as highlighted by Duff, is inherently complex and diverse.<sup>205</sup> This complexity is reflected in the varied goals of the testing cohort, highlighting the need for AI tools to be versatile and robust to address this diversity. (Fig.14)

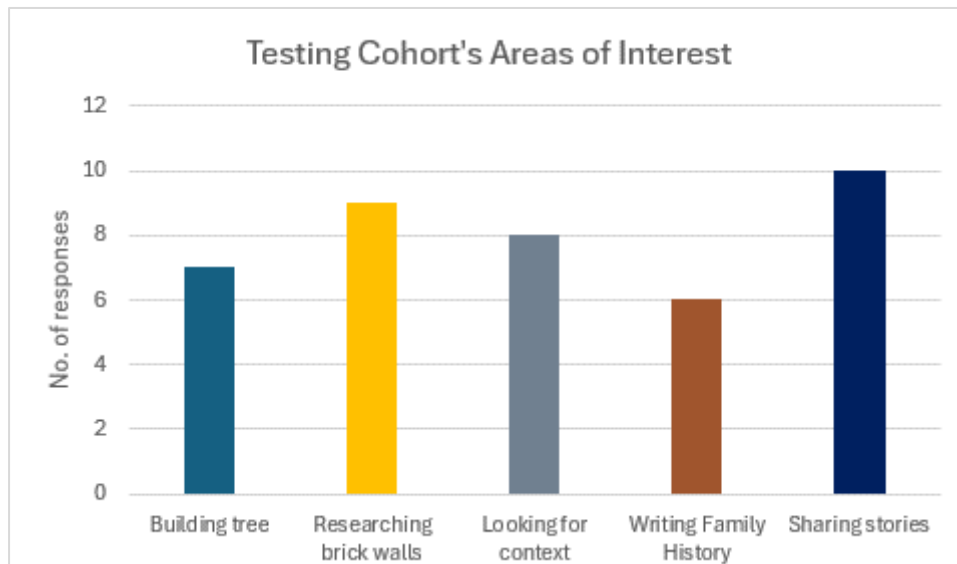


Figure 14 The testing cohort's diverse genealogical objectives.

To meet these varied needs, generative-AI must be able to organise, interpret and extract insights from data.<sup>206</sup> Using Little's categories as a framework, the CAs evaluated how generative-AI integrated into genealogical workflows.<sup>207</sup> These categories were used to discuss the research findings and provide an evaluation of the findings. (Table.5) Appendix B contains the raw results from the CAs, while Appendix C contains the processed data extracted from the CAs.

Category	Task Description	Capability Analyses
<b>Summarising</b>	Analysing data	CA4, CA6, CA7
	Summarising	CA5
<b>Extraction</b>	Data Extraction	CA2, CA3, CA5
<b>Translation</b>	Translation	CA3, CA5
<b>Generation</b>	Brainstorming	CA1
	Creating a family tree from data	CA5, CA7
	Creating a hypothesis	CA4, CA6
	Developing plans	CA8, CA7
	Explaining	CA4, CA7
	Providing context	CA3, CA4
	Writing a narrative	CA1, CA2

Table 5 Overview of generative-AI tasks by category.



## Summarising

Analysis of summarising results revealed respondents were impressed with Generative-AI's ability to accurately condense large texts into concise summaries, resulting in a high degree of usefulness (mean 3.9/5). (Figs.15-16)

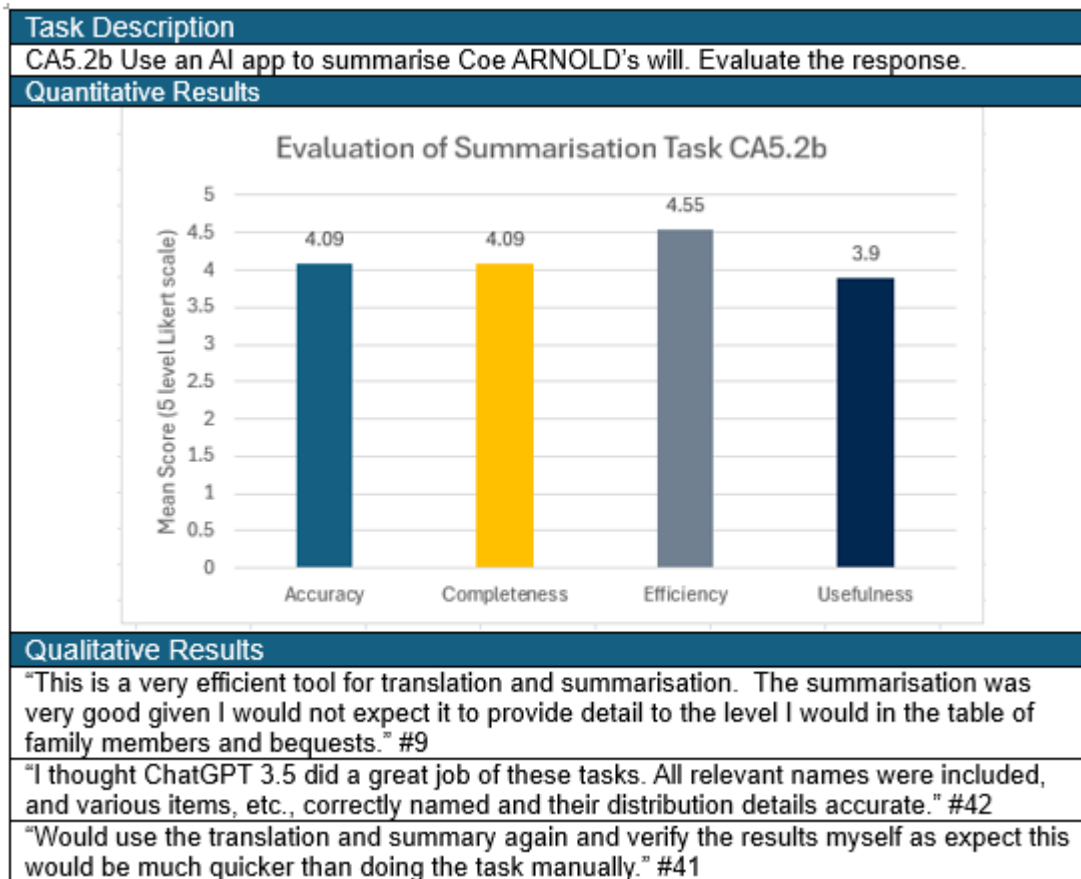


Figure 15 Evaluation of CA5.2 summarisation task.

## Summary of the Will of Coe Arnold (1676-1721)

---

**Name:** Coe Arnold

**Location:** Lowestoft, Suffolk

**Occupation:** Beer Brewer

**Date of Will:** March 18, 1720

### Key Points:

#### 1. Spiritual and Bodily Commendation:

- Commends his soul to God and his body to be buried by his Executors.

#### 2. Bequests and Devises:

- **To Brother Thomas Arnold:** All fish houses, lands, and hereditaments recently purchased from Mr. Robt. and Mrs. V.
- **To Son Matthew Arnold:** All residential houses, stables, outhouses, brewhouse, brewing utensils, tun cart, horses, and additional lands including those bought from Mrs. Munds and Thos. Arnold.
- **To Daughter Susanna Arnold:** Part of his house occupied by Henry Williams, Mount house, middle yard, and shared access with brother John.
- **To Daughter Elizabeth Arnold:** Barns, Almshouses, land, and adjoining yards. Also, £30 to be invested for her.
- **To Daughter Thomason Arnold:** £60 to be invested until she turns 21. If she dies before then, the amount is to be divided among surviving siblings.
- **To Son John Arnold:** Fish houses, malthouses, and associated yards and utensils. Also, part of the house now occupied by Robt. Holmes and other specific parts of the property.
- **To Daughter Mary Arnold:** Included in the contingency plan for minors and the distribution of household goods.

#### 3. Contingency Plan for Minors:

- If any child dies before reaching 21, their share is to be divided among surviving siblings. Executors are authorized to sell properties to provide for this distribution.

#### 4. Stock and Goods:

- All stock to be sold within a year of his death. Proceeds, along with other money and debts, to be used to pay debts, funeral expenses, legacies, and then divided among all children.

#### 5. Executors:

- Appoints brother Matthew Arnold of Lowestoft and cousin William Arnold of Ringfield as Executors, granting them £5 each for their efforts.

**Signed:** Coe Arnold

Figure 16 AI-generated summary of Coe ARNOLD's will. (Transcript in Appendix A CA5) <sup>208</sup> <sup>209</sup>

Generative-AI's summarisation ability offers significant speed benefits. For example, CA5.2b summarised a multi-page will in 0.88 minutes (mean). Including the additional 5 minutes (mean) required for validation, generative-AI proved significantly faster than the mean human time of 26 minutes. (Fig.17) Qualitative feedback highlighted the importance of validating results to ensure accuracy. These time savings were consistent across all tasks.

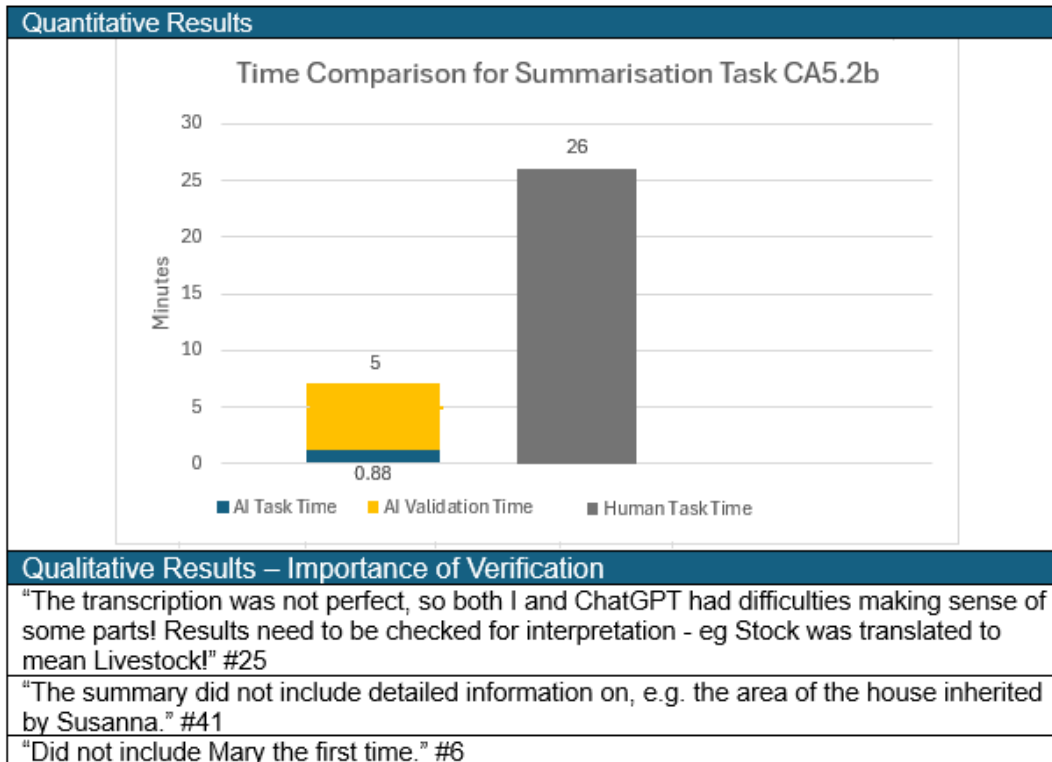


Figure 17 Comparison of the time taken for generative-AI and manual summary.

Efficiently summarising large amounts of unstructured data is helpful for many genealogical research activities like summarising reports and articles, consolidating land records or creating research summaries.

## Extraction

Generative-AI's information extraction capabilities were tested on various unstructured texts, including an obituary, a foreign birth record and a will. It achieved a mean accuracy score of 4/5 across four tasks, reliably identifying and extracting information swiftly and accurately.

(Fig.18)

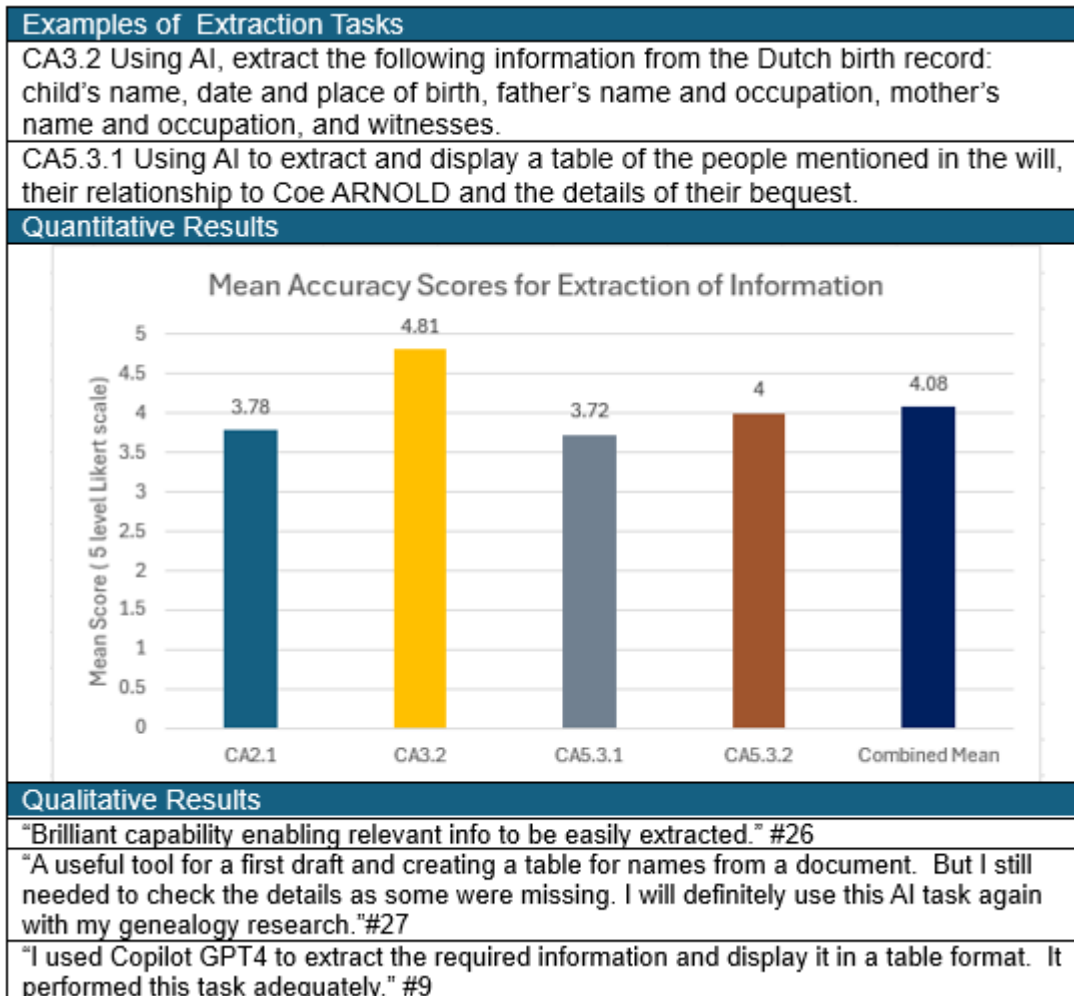


Figure 18 Evaluation of extraction accuracy.

The mean time taken for CA5.3.2, the most detailed extraction task, was 1.56 min plus 5.4 minutes for validation, a substantial timesaving compared to the 22 minutes (mean) for manual extraction. (Table.6)

Mean Time in Minutes	AI-extraction	Validation of Output	Manual Extraction
CA5.2.3 – Extract a list of the properties mentioned in Coe ARNOLD's will, the inheritor, and the supporting evidence in the will.	1.56	5.4	22

*Table 6 Comparison times for CA5.2.3 - Extraction*

This is particularly useful when information needs to be extracted in multiple ways, e.g. by property as well as person. See Fig19 for an extraction example.

## AI Extraction of Inheritance by Person

Name	Relationship to Coe Arnold	Bequest	Supporting Evidence
Thomas Arnold	Brother	All fish houses, lands, and hereditaments recently purchased from Mr. Robt. and Mrs. V	"I give, devise and bequeath unto my loving brother Thomas Arnold and his heirs forever all those fish houses, lands and hereditaments whatsoever I lately purchased from Mr. Robt. and Mrs.V"
Matthew Arnold	Son	All houses where Coe dwelt, the stables, outhouses, brewhouse, brewing utensils, tun cart, four horses, freehold Sun house Chamber Yard, barns bought of Mrs. Munds	"I give, devise and bequeath unto my son Matthew and his heirs forever all ye houses wherein I dwelt..."
Susanna Arnold	Daughter	Part of house occupied by Henry Williams, Mount house and middle yard	"I give devise and bequeath unto my daughter Susanna Arnold and her heirs forever all that part of my house in ye occupation of my brother Henry Williams..."
John Arnold	Son	Fish houses, Mr. Hibors Malthouses, house occupied by Robt. Holmes, part of house occupied by William, upper yard	"I give, devise and bequeath unto my son John Arnold and his heirs forever all those my Fish Houses, lately Mr. Hibors Malthouses..."

## AI Extraction of Inheritance by Property

Property	Inherited By	Relationship to Coe Arnold	Supporting Evidence
Fish houses, lands, and hereditaments	Thomas Arnold	Brother	"I give, devise and bequeath unto my loving brother Thomas Arnold and his heirs forever all those fish houses, lands and hereditaments whatsoever I lately purchased from Mr. Robt. and Mrs.V"
Houses where Coe dwelt, stables, outhouses, brewhouse, brewing utensils, tun cart, four horses, freehold Sun house Chamber Yard, barns bought of Mrs. Munds	Matthew Arnold	Son	"I give, devise and bequeath unto my son Matthew and his heirs forever all ye houses wherein I dwelt, the Stables and Outhouses thereunto belonging together, the Brewhouse, coppers and all ye Brewing Utensils whatsoever, the tun cart and four horses..."
Part of house occupied by Henry Williams, Mount house and middle yard	Susanna Arnold	Daughter	"I give devise and bequeath unto my daughter Susanna Arnold and her heirs forever all that part of my house in ye occupation of my brother Henry Williams as lyeth by the north of ye Balk in or about ye middle of ye Great Room, together with Mount house and middle yard..."

**Task description CA5.3.2:** Use AI to extract and display a table of the people mentioned in the will, their relationship with Coe ARNOLD and the details of the bequest. Then extract and display the table of the properties mentioned in the will, who inherited them and their relationship to Coe ARNOLD.

Figure 19 Excerpts illustrating AI extraction of information from Coe ARNOLD's will. (Transcript in Appendix A CA5). <sup>210 211 212</sup>

## Translation

Generative-AI can translate many languages, helping make foreign archival records accessible. It can also rephrase archaic texts into contemporary language, enabling quick assessments of documents for value and relevance. The analysis revealed the generative-AI translation delivered contextually accurate, valid translations for CA3.2 and CA5.2. The efficiency rating for the tasks was high, 4.6/5, making this feature a valuable addition to the research toolbox. (Fig.20)

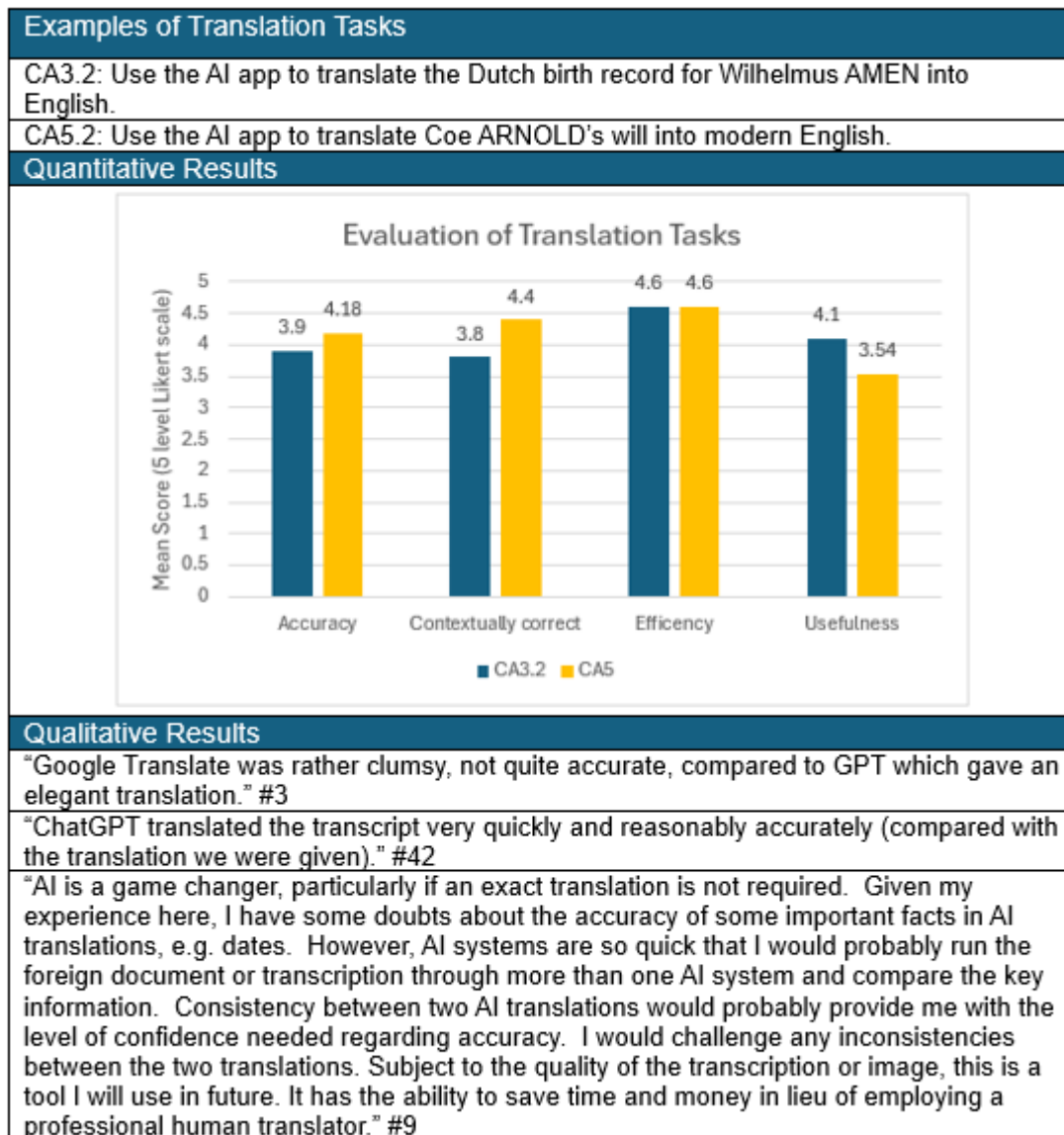


Figure 20 Evaluation of translation tasks.

Fig. 21 compares a verified human translation of the CA3 Netherlands birth record with the generative-AI version. While the layouts differ, the generative-AI version captured all the names, dates and places, allowing genealogists to extract the relevant information quickly (mean time 103 seconds). However, Hoitink advises that translations, notably less common languages, should be treated cautiously, as generative-AI can produce plausible but inaccurate information.<sup>213</sup> A human translator is still the best option if 100% accuracy is required.



# FamilySearch Translation

## Translation of Birth Registration

Wilhelmus Johannes Franciscus Henricus Amen, Geboren no. 3/38, Burgerlijke Stand, Amsterdam, Noord Holland, Nederlands.

On this day the twenty-second of March Eighteen Hundred and Forty-One, is before the undersigned Civil Officer from the Civil Registry Office of Amsterdam, appears: Henricus Amen of the occupation wagon maker aged thirty-nine Years, living at Guldenhandsteeg Canton N<sup>o</sup>. 124, who has declared that on this day at six o'clock a.m., in the house, standing here is born a child of the male sex, from Hendrica Beernink of the occupation none living here his wife whose child shall be named Wilhelmus Johannes Franciscus Henricus of which we declare this Act has been made in the presence of Johan Michael Giersberg of the occupation wagon maker aged thirty Years, living at Nuijenstraat n1058. and of Hendricus Gerardus de Jong of the occupation seaman aged forty-two Years, living at Dijkstraat and this Act through us, along with [ ] testify, after reading is signed.

H. Amen [signature]  
J. M. Giersberg [signature]        }        Dentz [signature]  
H: G: de Jong [signature]

# ChatGPT-4o Translation

## Translation of Birth Registration

Wilhelmus Johannes Franciscus Henricus Amen, Born no. 3/38, Civil Registry, Amsterdam, Noord Holland, Netherlands.

On the twenty-second of March, eighteen hundred forty-one, appeared before me, the undersigned official of the Civil Registry of the city of Amsterdam:

**Henricus Amen**, occupation: wagonmaker, age thirty-nine years, residing at Guldenhandsteeg, Canton No. 124, who declared that on this day, at six o'clock in the morning, in the house located here, a child of the male gender was born, from **Hendrica Beernink** occupation: none, residing here, his lawful wife.

This child was given the names:  
**Wilhelmus Johannes Franciscus Henricus**

This declaration was made in the presence of Johan Michael Giersberg, occupation: wagonmaker, age thirty years, residing at Nuijenstraat No. 1058, and Hendricus Gerardus de Jong, occupation: sailor, age forty-two years, residing at Dijkstraat, and this record has been signed by us, as well as the witnesses, after it was read aloud.

**\*\*Signatures:\*\***  
H. Amen [signature]  
J. M. Giersberg [signature]        } **\*\*Dentz\*\*** [signature]  
H. G. de Jong [signature]

Figure 21 Comparison of FamilySearch manual translation with generative-AI translation. <sup>214 215 216</sup>

## Generation

AI's generation of narrative content from existing data is highly useful. It can create narrative reports from lists of names, dates and relationships or timelines, brainstorm ideas for presentations, suggest alternative spellings of surnames, or provide context. (Fig.22) However, the accuracy and quality of the generated texts depend on the quality of the prompt and information provided, emphasising the need for careful verification. <sup>217 218</sup> Appendix D has more information on prompt creation.

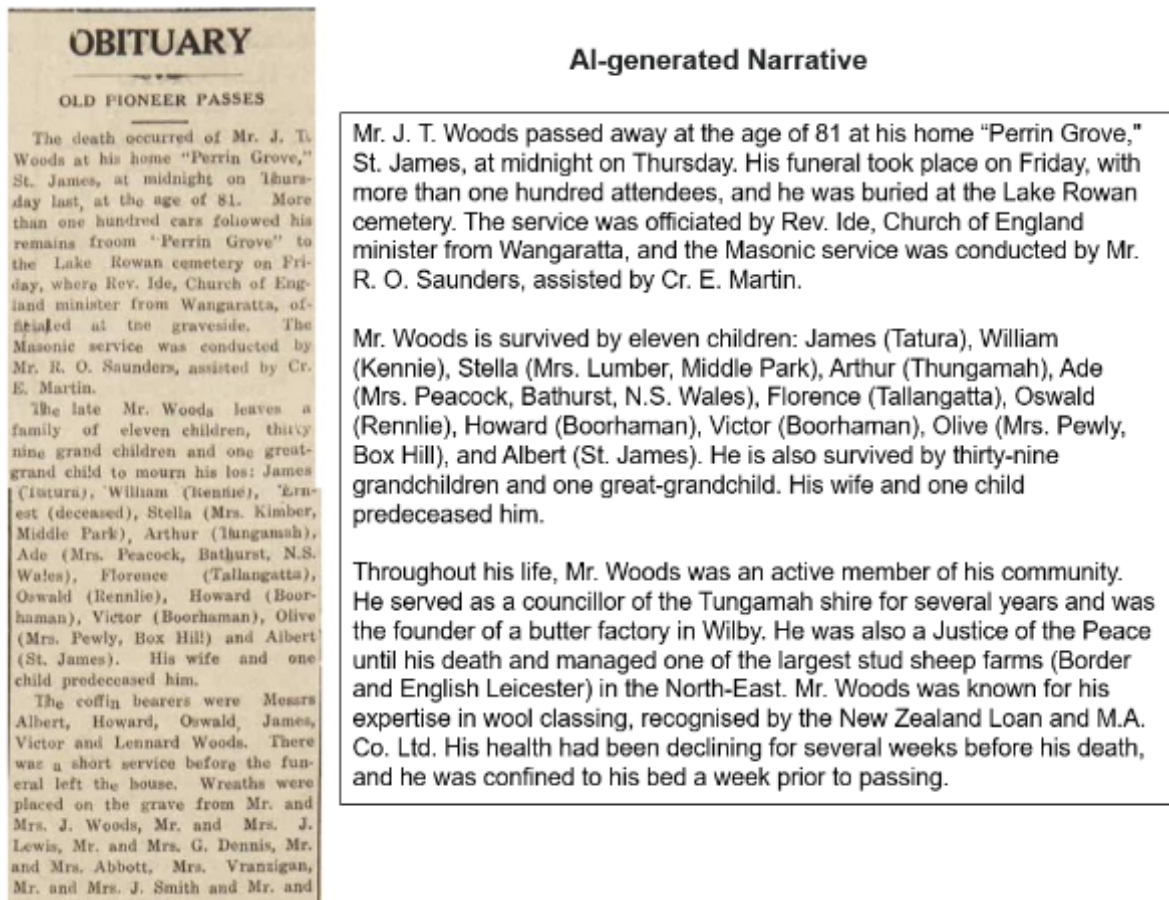


Figure 22 Example of AI-generated narrative from an obituary – CA2.3. <sup>219 220</sup>

Evaluation of generation tasks across a range of CAs revealed respondents found generative-AI's performance in these styles of task to be very accurate, efficient and useful. (Fig.23). Their feedback illustrated the benefits of generative-AI as a collaborative tool to augment their skills, provide different perspectives on data and create valuable narratives.

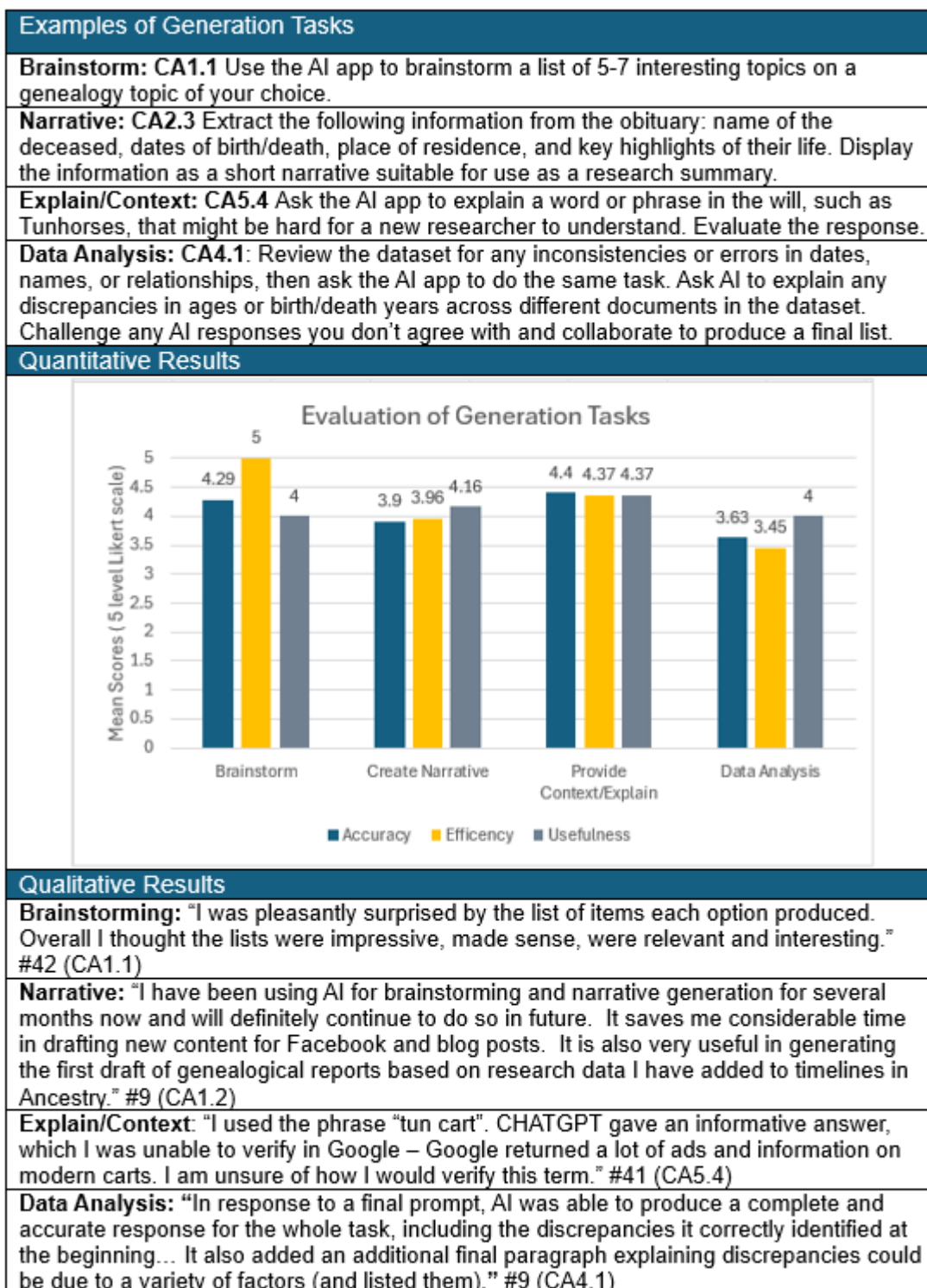


Figure 23 Evaluation of generation tasks.

## Awareness of Generative-AI for Genealogy Research

Although the cohort had expertise in genealogy and computer skills, the preliminary survey revealed that, like many genealogists, they had a limited engagement with AI overall and even less experience with generative-AI before participating in this research. (Fig.24)

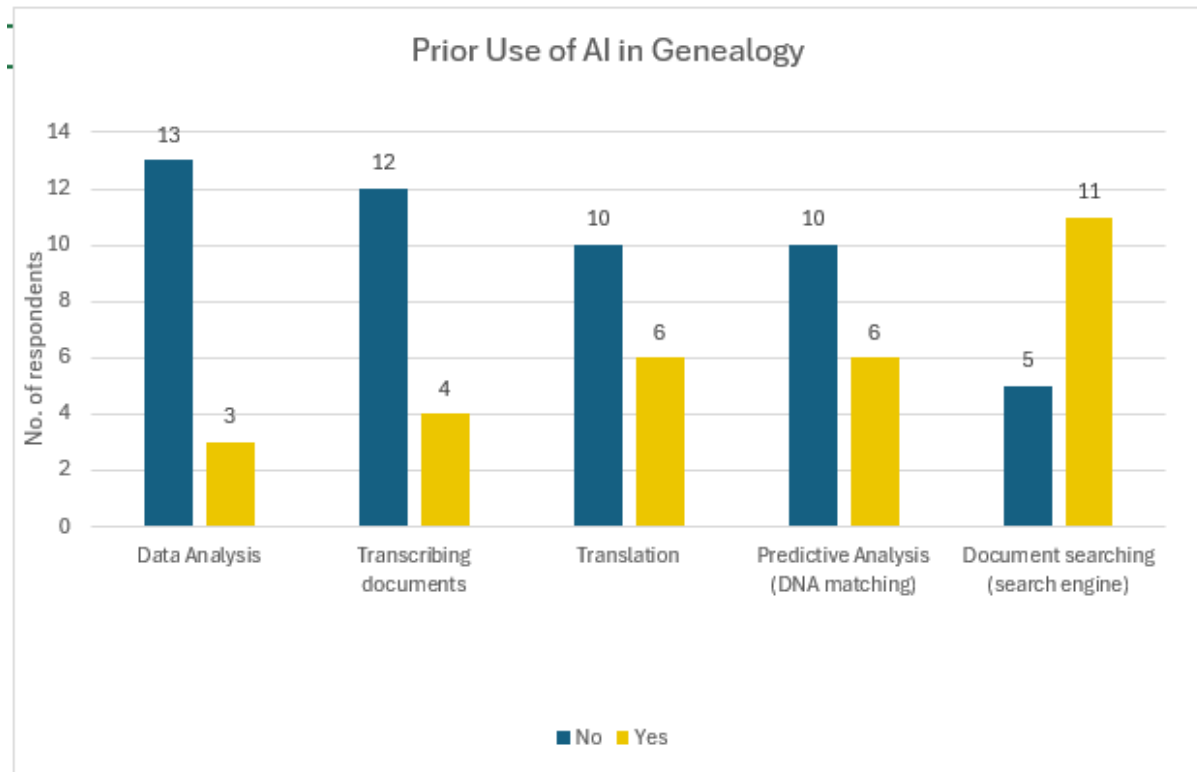


Figure 24 Preliminary Survey results on prior use of AI for genealogy workflows.

However, anecdotal evidence shows that genealogists are increasingly adopting generative AI for their research, albeit cautiously, due to its limitations. Evidence of this trend includes the growing membership of the *Genealogy and Artificial Intelligence* Facebook group (6.7K in April 2024 and 7K in June 2024), an increase in generative-AI presentations at RootsTech from 2023 to 2024, and a Top 10 ranking for generative-AI presentations by Leeds and McEntee among Legacy webinars in March 2024.<sup>221 222 223</sup>

This growing integration of generative-AI into genealogy workflows aligns with Rogers' theory, which posits that adopting new technologies follows a predictable pattern with more community members using it over time.<sup>224</sup>

## Benefits of Generative-AI in Genealogy

Genealogy relies on factual integrity and accuracy achieved through careful analysis of records. Generative AI, known for its word-centric algorithms and a tendency to hallucinate, might appear ill-suited to such a precise field. However, analysis of the CA data and the debriefing survey revealed 73% of the testing cohort found generative-AI valuable for enhancing the accuracy and depth of their research. (Table.7)

Qualitative Results – Benefits of Generative AI (selected from all CAs)
"Overall, though, the results showed that it can be very useful for quick and easy assistance with translating, collating and summarising information from long and unwieldy texts to produce useful research aids in various formats. However, this needs to be done in stages, carefully checking each ChatGPT response for errors and omissions and rephrasing the prompt, if necessary, to ensure the validity of the final result." #11 (CA5)
"Very useful as a collaborative tool - AI picked up things I missed, and I had to prompt to get completely accurate information, but we worked well together. Yes, I would use in the future." #25 (CA2)
"As I had not used this before for Genealogy, I was amazed at the capability to expand and respond." #3 (CA1)
"It was good to have another perspective, the list ChatGPT came up with even though general was solid. It forced me to think more deeply about the task, which made me realise I'd made some assumptions that could easily be wrong. Even when I thought it went in the wrong direction, that was good as it forced me to think about why it was the wrong direction and that sparked other ideas in me. I like its bottomless patience for working on a task." #16 (CA4)

Table 7 Sample of qualitative feedback illustrating the positive impact of generative AI.

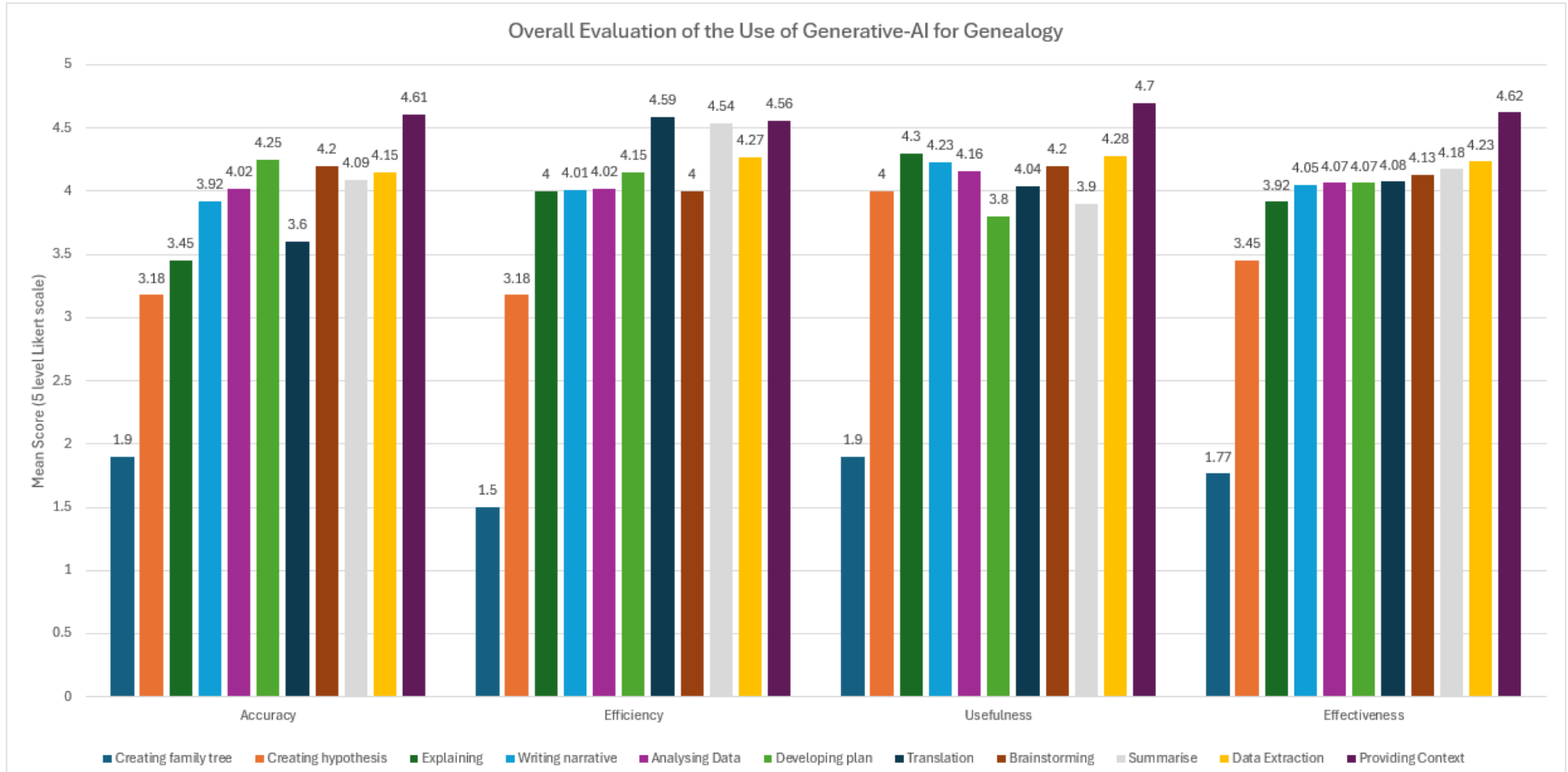


Figure 25 Quantitative analysis of generative AI's performance across all task categories.

The accuracy, efficiency, usefulness and overall effectiveness of generative-AI across all the CA tasks are demonstrated in Fig.25 It reveals that with proper oversight, generative-AI significantly enhances genealogical research.

High effectiveness scores in Data Extraction and Providing Context illustrate that generative-AI is highly capable of managing genealogical data, extracting relevant information with a precision comparable to manual methods and providing context to assist accurate data interpretation. Tasks like Translation, Summarising Information, and Research Planning demonstrate AI's assistance in translating documents, distilling records into summaries and strategising research approaches. These findings highlight generative-AI's ability to augment, not replace, genealogists' skills. These findings are consistent with Dell'Acqua and Hörnemalm's research, which shows that performance can improve up to 43% when ChatGPT is used collaboratively.<sup>225 226</sup>

Potential relationships between Accuracy, Efficiency and Usefulness were determined using Pearson's Correlation Coefficient ( $r$ ), which measures the strength and direction of a linear relationship between two variables.<sup>227</sup> A rating of 0 indicates no relationship, 1 indicates a strong positive relationship, and -1 indicates a strong negative relationship. The analysis revealed a strong positive relationship between the accuracy, efficiency, and usefulness of generative-AI, showing the more accurate the AI-generated data, the more efficient and valuable it becomes for genealogists. (Table.8). These correlations reflect a generally favourable view of AI's capabilities for genealogy research, emphasising the practical benefits it can offer.

Correlation between Accuracy, Efficiency and Usefulness ( $r$ )	
Accuracy to Efficiency	0.86
Accuracy to Usefulness	0.80
Efficiency to Usefulness	0.85

*Table 8 Relationship between accuracy, efficiency and usefulness.*

The Effectiveness scores also indicate that the cohort used generative-AI in ways that capitalised on its strengths, while the high Efficiency scores demonstrated that AI expedites time-consuming aspects of data management, as illustrated earlier in Fig.25. The Accuracy results indicate the respondents' genealogy expertise was used to evaluate AI-generated results, minimising inaccuracies. These results indicate the respondents used AI as a powerful tool, not a standalone solution.

These findings align with literature review insights, suggesting that if generative-AI's challenges are recognised and managed, its application in genealogy can be highly beneficial.<sup>228 229</sup>

### Limitations of Generative AI in Genealogy

Emerging technologies like generative-AI pose risks as they have uncertain and wide-ranging capabilities that change at an exponential rate, necessitating a cautious approach to their use in genealogy.<sup>230 231 232 233</sup> The findings confirm these insights, revealing that while generative-AI is capable, it is not always the optimal tool for every task. (Table.9) This aligns with Mollick and Dell'Acqua's concept of a "jagged frontier", which illustrates the uneven advancement of AI capabilities across different tasks.<sup>234 235</sup>

Qualitative Results – Limitations of Generative AI (selected from all CAs)
"More time, patience and additional prompts would be needed to get the required level of detail and accuracy. Based on this task alone, it is borderline whether I would use the tool for data extraction given the time needed for validation, multiple re-prompting and re-validation." #9 (CA3)
"I would definitely not rely on ChatGPT for data analysis as it missed several discrepancies. However, if I had a large dataset, I might use it to check if there were discrepancies that I had missed. I would probably not use it for small datasets as it was quite time-consuming to validate responses that were erroneous." #41(CA4)
"Did AI tell me anything new? Not really. An LLM is only as good as its database. Is it capable of making the kind of non-linear jumps to a new direction that a human brain is capable of? I don't think so." #18 (CA2)

*Table 9 Sample of qualitative feedback illustrating some of generative-AI's limitations.*

The primary concern is generative-AI's hallucinations, with 81.8% of respondents in the debriefing survey worried about introducing errors into their research. Prompt creation strategies advocated by Little and Mollick significantly reduce hallucinations but are not foolproof.<sup>236 237</sup> Appendix.D includes examples of narratives created with and without their prompt engineering recommendations. While seemingly credible, AI's confident and articulate responses can be misleading. Genealogists may struggle to distinguish between factual information and AI-generated fabrications, increasing the burden of verification. Respondent feedback indicated that the verification process is tedious but critical in deciding whether to utilise generative-AI.(Table.10)



Qualitative Results (selected from all CAs)
"I also hate having to check someone else's work – it's time-consuming and boring, and I feel more error-prone because my brain is refusing to do this." #16 (CA 4)
"I feel I can more readily determine if the answer is accurate and in context judging by its source than by asking the same thing of AI. Given AI systems' propensity to make things up, I feel I would nearly always have to validate the answer by checking another system. This would be inefficient and defeat the purpose of using AI." #11 (CA5)

Table 10 A sample of qualitative responses illustrating the tedious nature of verification.

The findings also revealed limitations in generative-AI's performance for some complex tasks like CA5.3.2. (Fig.26) Although mean scores were high, indicating a positive perception of generative-AI's capabilities, analysis revealed individual scores varied significantly, reflecting respondents' differing opinions on performance.

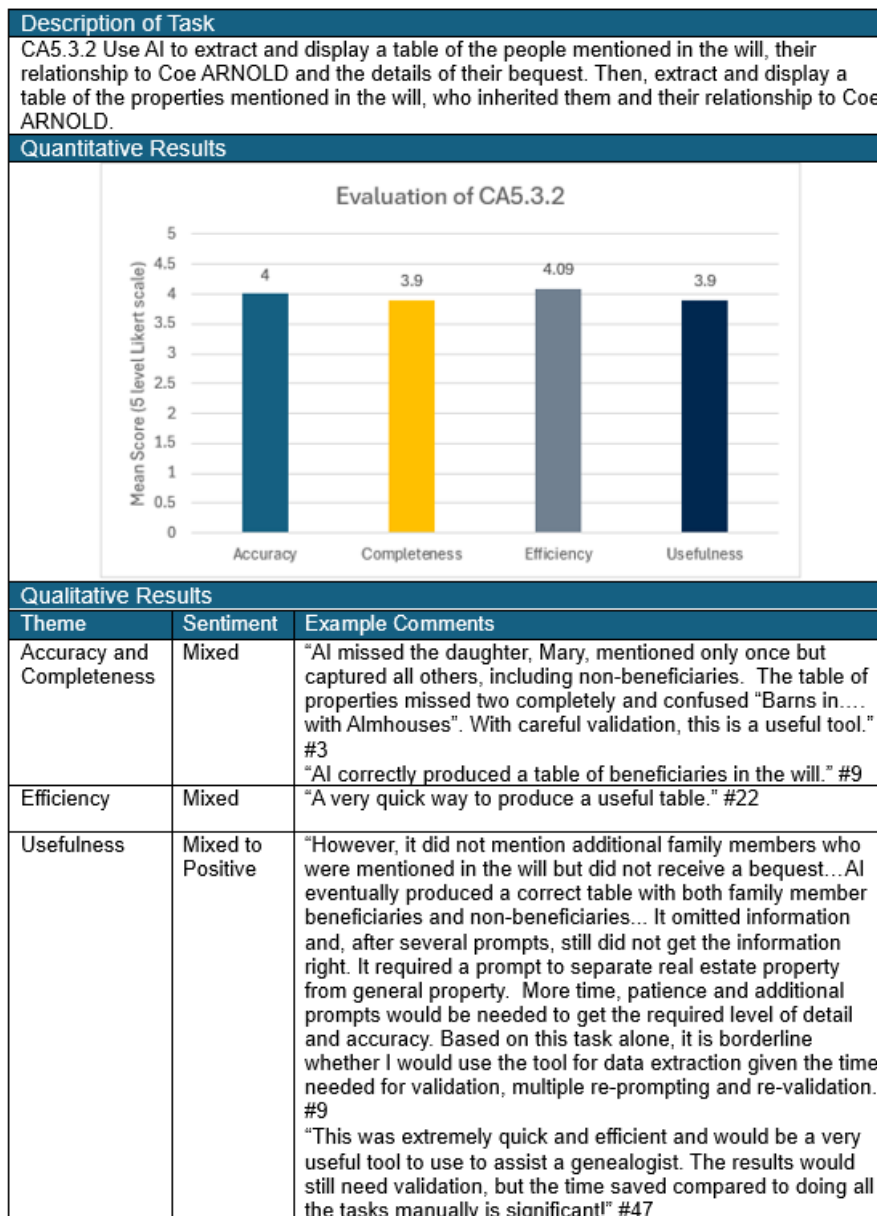


Figure 26 Analysis of the results for CA6.3.2 illustrating the high variability of responses.

Analysing the quantitative scores using a box and whisker plot reveals the spread and central tendency of the data, highlighting areas of agreement and disagreement among respondents.

1. **Median Score:** The median score in each category (X) is approximately 4, indicating that half the respondents rated generative-AI's performance above this level.
2. **Interquartile Range (IQR):** In the first three categories, the IQR of scores is between 4-5. This range, shown by the box's bottom and top, represents the middle 50% of responses, indicating a consensus in these categories
3. **Whiskers:** The whiskers of the plot for these categories extend from 3-5, covering the 10<sup>th</sup>-90<sup>th</sup> percentiles, indicating the majority of respondents rated generative-AI's performance between these values with a few outliers (dots) outside this range.

This clustering of scores within a narrow range suggests close agreement on generative-AI's performance in most areas. However, the Usefulness category shows greater variation, with scores spreading more widely. This suggests mixed experiences, as some respondents found generative-AI very useful, while others rated it much lower, reflecting different expectations or experiences. This is consistent with comments in the qualitative data.

(Fig.27)

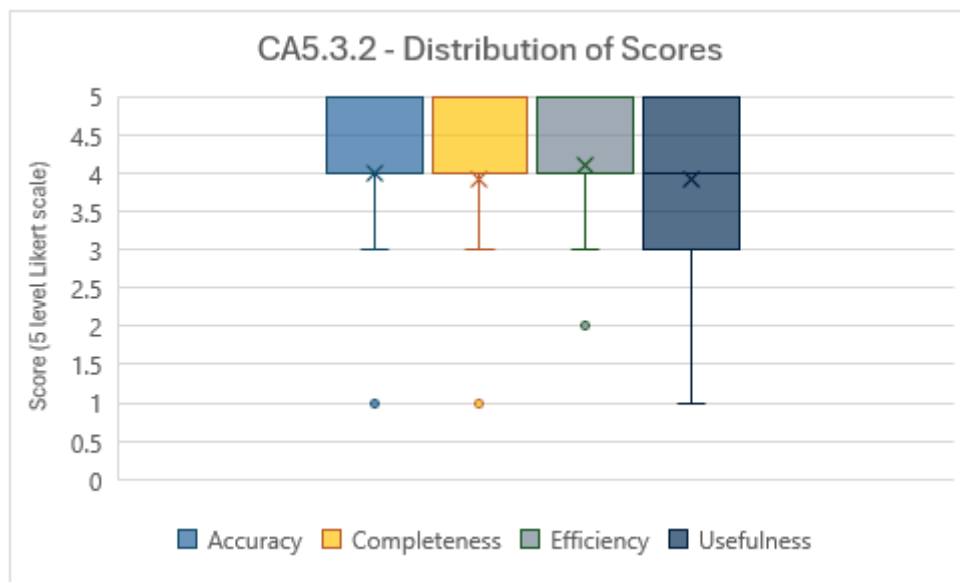


Figure 27 Box and Whisker chart showing the distribution of scores for CA5.3.2.

The greater variation in the Usefulness category can be attributed to its subjective nature. Individual perceptions of usefulness are influenced by personal familiarity with the task and prior experience with generative-AI, leading to diverse evaluations. As discussed earlier, there is also a strong correlation between the other categories and Usefulness ( $r$  ranges

from 0.87 to 0.88), with a weaker result in one of the other areas impacting the perceived usefulness of the result.

The presence of outliers indicates some respondents rated these metrics significantly lower than the majority. This could be because experienced individuals may find it more efficient to perform tasks without AI assistance or a lack of knowledge in a particular area, making it difficult to verify a response, as suggested by the beginner-level genealogist scores.<sup>238</sup> (Table.11)

Quantitative Feedback	Accuracy	Completeness	Efficiency	Usefulness
It would probably still be quicker to edit CHATGPT's output than start from scratch as CHATGPT's table copied	1	1	2	3

Table 11 Beginner-level scores for CA5.3.2.

However, the statistical correlation between genealogy skill levels and task scores was very weak ( $r=0.014$ ), suggesting skill level did not significantly impact evaluations. The small sample size ( $n=11$ ) limits this analysis, making it challenging to detect significant relationships. Further investigation with a larger dataset would be needed for a more robust conclusion.

Generative AI's non-deterministic nature also introduces a level of unpredictability, with identical queries yielding different responses at different times. This unpredictability could limit the use of generative-AI in genealogical research, where data reliability is essential. All generative-AI content must be carefully assessed and interpreted to ensure its accuracy. However, genealogists are accustomed to dealing with inaccurate data, and the qualitative data indicated that the verification process did not significantly detract from its usefulness.

The qualitative data mentioned earlier in Fig.26, also indicated the quality of prompts heavily influences the effectiveness of generative-AI. Respondents noticed that well-crafted prompts were crucial for accurate and complete responses, efficient data extraction, and overall usefulness. Expertise in a skill enables the creation of more contextually appropriate prompts, enhancing response quality as well as better evaluation of output.<sup>239 240</sup> While crafting effective prompts can be time-consuming, their value increases with repeated tasks, as prompts can be refined and reused for consistent results. This analysis highlights the need for further exploration into specific aspects of prompt creation to enhance the effectiveness and user satisfaction of generative-AI.

Complex tasks involving ambiguous data also posed significant challenges for generative-AI. These “jagged frontier” tasks, often cause the AI to generate inaccurate information.<sup>241</sup> For example, CA5.5 required generative-AI to construct a family tree based on a will, a common but complex genealogical task.

Current AI systems struggle with this task due to their limited understanding of genealogical relationships. The results show generative-AI performed poorly for Completeness, Usefulness and Efficiency, although some respondents felt the family tree might be useful. (Fig.28) Figure.29 showcases examples of the AI's responses and illustrates the performance difference between ChatGPT-3.5 and the newer ChatGPT-4o. The marked improvement within three months strongly suggests that generative AI's capabilities in this area will continue to advance.

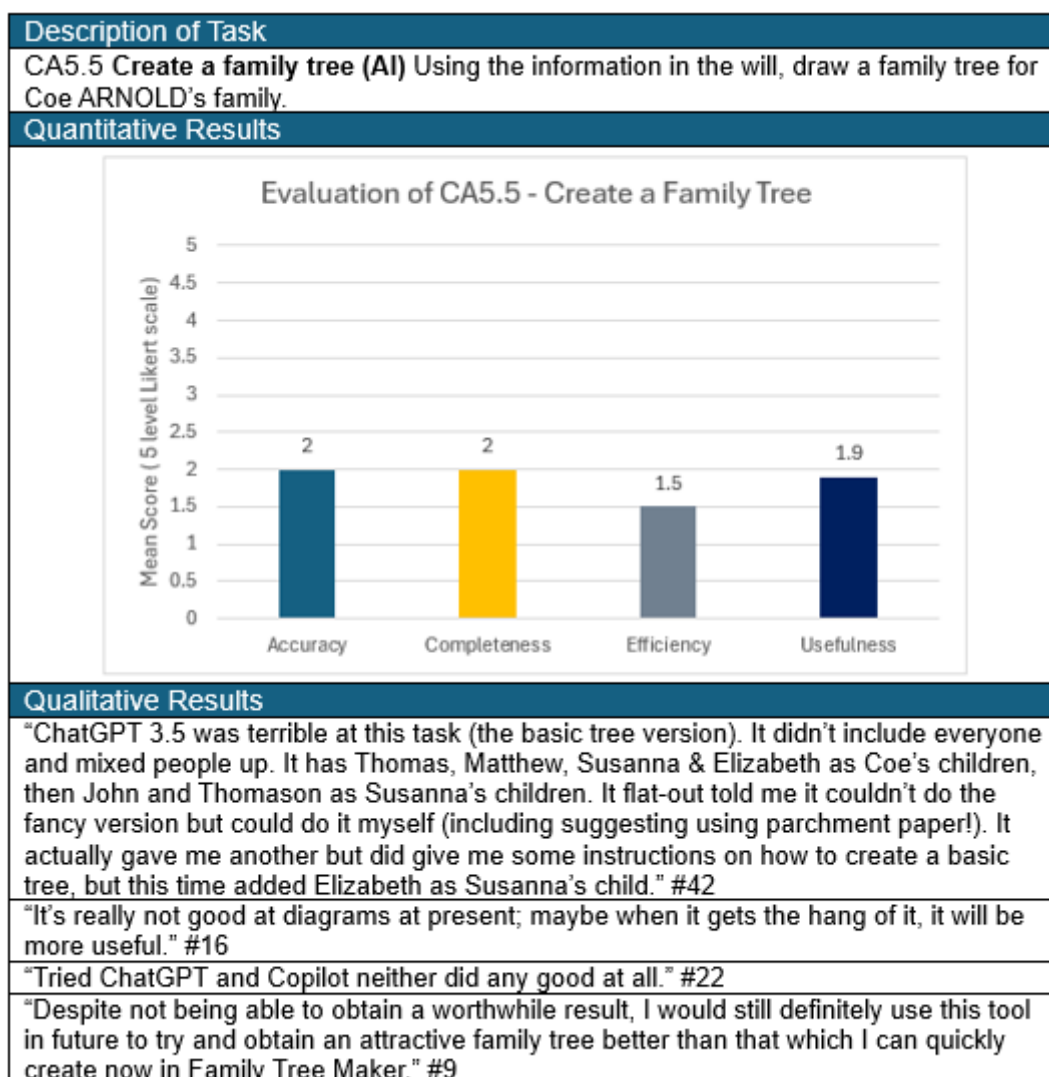


Figure 28 Analysis of generative-AI's performance in “jagged frontier” tasks.

### Extract from Transcript of Coe ARNOLD's Will

Item. I give, devise and bequeath to my daughter THOMASON ARNOLD ye sum of Sixty pounds to be raised out of my estate and put out to interest for her until she attains unto ye age of twenty-one years and then to be paid to her and if she dies before to be equally divided between her brothers and sisters that survive. It is my mind and will also is ye aforesaid Thirty pounds given to my daughter Elizabeth and all ye aforesaid houses given unto my sons Matthew and John and my daughters Susanna, Mary and Elizabeth should be, in case they or any of them should die before they attain unto ye age of twenty-one years and in case ye aforesaid heirs shall neglect or refuse to pay unto my surviving children their respective share of ye aforesaid worth or value of ye aforesaid houses or legacy before given to my children shall happen to depart this life before they attain ye aforesaid age of twenty-one years, the worth or value such houses is to be sold by ...

#### Extraction of Data (ChatGPT-4o)

Matthew Arnold	Son
Susanna Arnold	Daughter
John Arnold	Son
Elizabeth Arnold	Daughter
Thomason Arnold	Daughter
Mary Arnold	Daughter
Henry Williams	Brother-in-law (brother of Coe's wife)
Matthew Arnold	Father (deceased)
Mrs. Munds	Acquaintance (widow)
William Arnold	Brother

### AI-generated Family Tree

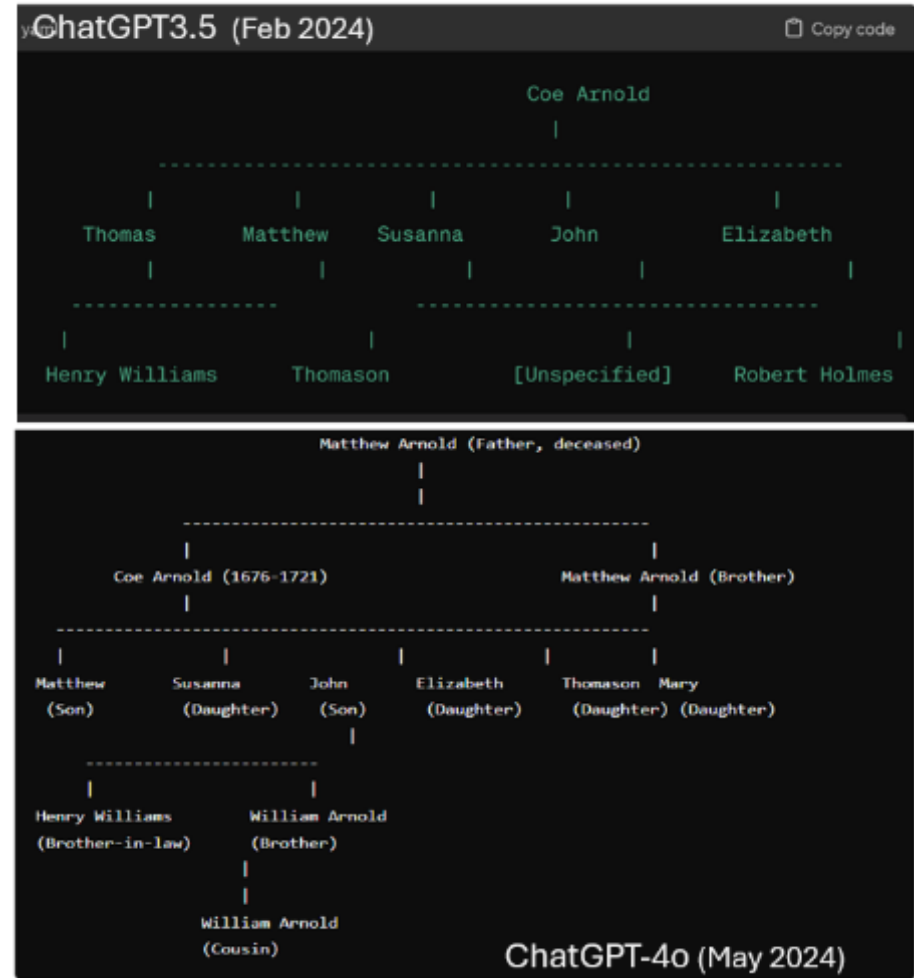


Figure 29 Illustration of generative-AI ability to create a family tree. 242 243 244 245

## **Other Considerations**

An important consideration in the adoption of new technology is ensuring a supportive environment for users, particularly in the early phases of integration into workflows. The withdrawal of two participants after the introductory workshop, despite their high levels of interest indicated in the preliminary survey, suggests they were not ready to adopt the technology at this early stage. This aligns with Rogers' theory on the innovator/early adopter phases of innovation adoption.<sup>246</sup> As recommended in Lucy's research, the introduction of generative-AI tools should be accompanied by peer-support measures tailored to the different adopter categories to facilitate a smooth transition.<sup>247</sup> Additionally, there are significant ethical concerns surrounding the use of generative-AI, which will be explored in section 4.2.

## **Conclusion**

In keeping with the insights from the literature review, the analysis of the findings revealed that despite the limitations, generative AI can be effectively used for genealogy research, improving the speed and efficiency of many common activities. The best results were achieved when the respondents worked collaboratively with generative-AI, using its abilities to supplement their own skills. Although generative-AI's responses require validation, this is no different to any other genealogical source, and researchers can capitalise on AI's capabilities while maintaining the accuracy and integrity of their genealogical inquiries.

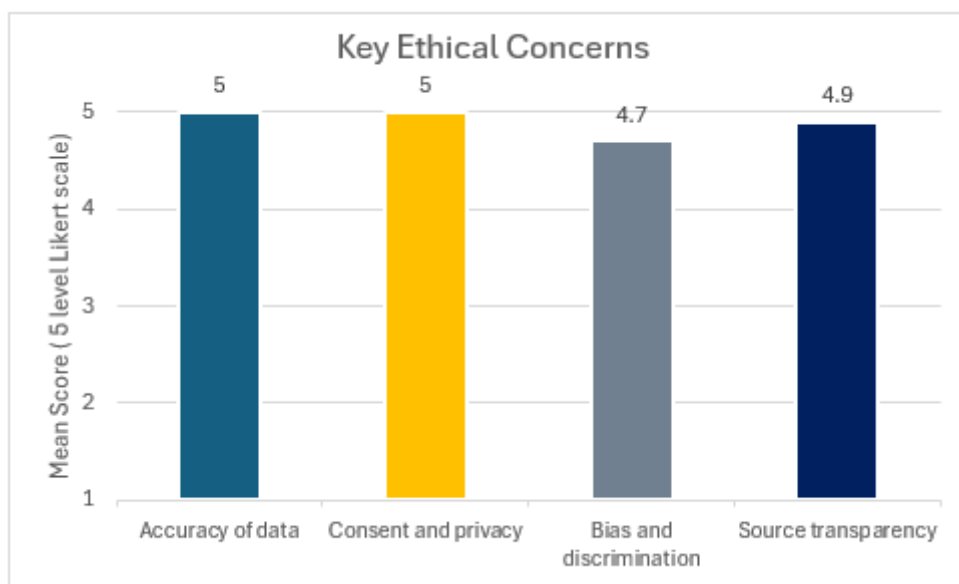
## 4.2 The Ethical Considerations of AI in Genealogy:

### ***RQ 3: What ethical considerations and safeguards are needed when using AI, and what risks arise from not implementing them?***

The literature review revealed that integrating generative-AI technologies into genealogy research presents critical ethical challenges.<sup>248 249 250</sup> Some of these challenges are being addressed at a global level; however, others need to be addressed by the genealogical community.<sup>251 252 253</sup> CA6 collected information on ethical concerns, ethical guidelines for AI use, and AI's ability to resolve ethical questions. This section merges insights from CA6 data with literature review findings to outline guidelines for the ethical integration of AI into genealogy.

#### **Key Ethical Considerations.**

Analysis of CA6 identified key ethical concerns from respondents regarding the use of AI, including data accuracy, the need for consent and robust privacy measures, the potential for bias and discrimination, and the importance of transparent sourcing. (Fig.30) These concerns align with those identified in the literature review, emphasising their universal relevance across various fields, not just genealogy.<sup>254 255 256</sup>



*Figure 30 Ethical concerns identified by testing cohort.*

There was a significant change in the respondents' ethical concerns about AI in genealogy over time. The rise in concern during CA6 (midway through data gathering), followed by a significant decrease at the end, highlights a peak in concerns that diminishes but doesn't return to preliminary levels by the end of the study. (Fig.31)

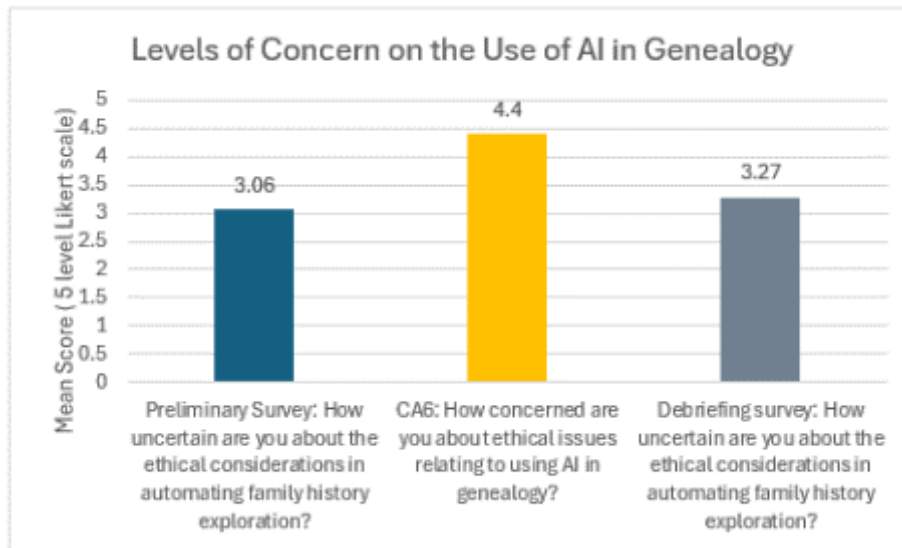


Figure 31 Evaluation of respondents' concern on the use of AI-assisted genealogy.

No statistically significant correlation was found between the level of uncertainty about ethical considerations in the preliminary and debriefing surveys and the concerns measured in CA6. The limited sample size suggests that any observed correlations should be interpreted with caution and that larger sample sizes are needed for more definitive conclusions.

Possible explanations for the observed differences include increased familiarity with generative-AI, leading to higher concern levels that diminished as respondents became more comfortable with the technology. Rogers' innovation-decision process can be used to understand these variations as part of the natural adoption process.<sup>257</sup> (Fig.32)

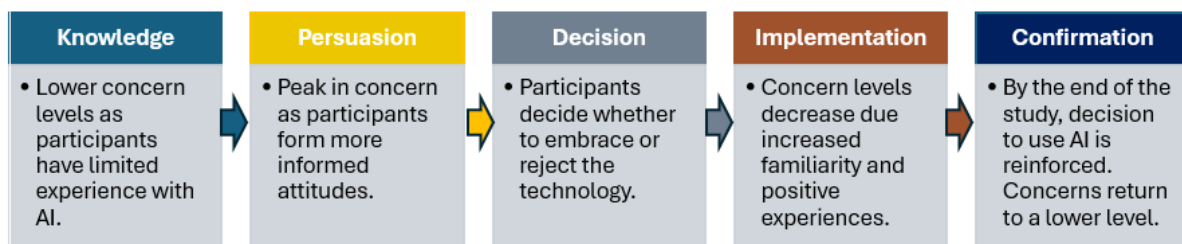


Figure 32 Application of Rogers' Innovation-Decision Process.<sup>258</sup>



While subtle differences in question-wording, such as “automation” versus “AI”, might have influenced responses, external factors, like increased awareness of ethical issues due to media reports or CA6’s detailed ethics discussion, could also play a role. Further investigation with more precisely worded questions and a larger sample size is needed to accurately pinpoint the factors contributing to these differences.

### Existing Genealogy Standards

Current genealogy standards emphasise the importance of verifiable, well-sourced, and unbiased data for accurate, valid research.<sup>259 260 261</sup> However, while a majority of respondents believe the standards proved an effective framework for using AI, opinions are notably divided. (Fig.33) There is no significant statistical relationship between the respondents' skill levels and their scores ( $r = 0.37$ ). The variation may stem from varied expectations and personal standards of adequate ethical guidance. This finding suggests there are areas where the standards need updating, as advocated by Bettinger.<sup>262 263</sup> Further research is needed to explore how the standards can be improved to better support ethical AI-assisted genealogy.



Figure 33 Evaluation of applicability of existing genealogical standards.

## Developing and Evaluating Better Frameworks

Respondents also evaluated a set of ethical guidelines created by generative-AI. The suggested guidelines align with existing standards but are more applicable to AI-assisted genealogy. (Table.12). These guidelines received high relevance (mean 4.5/5) and effectiveness ratings (mean 4.2/5), suggesting they provide a robust framework for ethical AI integration. Appendix.E contains the AI-generated framework.

Generative AI Framework for AI-assisted Genealogy Research	
1	Principle of Accuracy and Verification
2	Principle of Transparency
3	Principle of Privacy and Data Protection
4	Principle of Bias Mitigation
5	Principle of Ethical Use and Human Oversight
6	Principle of Intellectual Property
7	Principle of Continuing Education
8	Principle of Community Engagement
9	Principle of Reporting and Accountability

Table 12 Outline of AI-generated framework for ethical AI-assisted research.<sup>264</sup>

Respondents viewed AI as a tool complementing traditional methods, focusing on transparency and detailed documentation. Echoing literature review insights, they recommend treating AI outputs with a proactive but cautious approach, emphasising verification, human oversight and ethical standards.<sup>265 266 267</sup> Other suggestions included developing specific business or personal AI policies and addressing ethical concerns such as privacy and copyright. Their recommendations balance technological advancement with the integrity of genealogical research. (Fig.34)

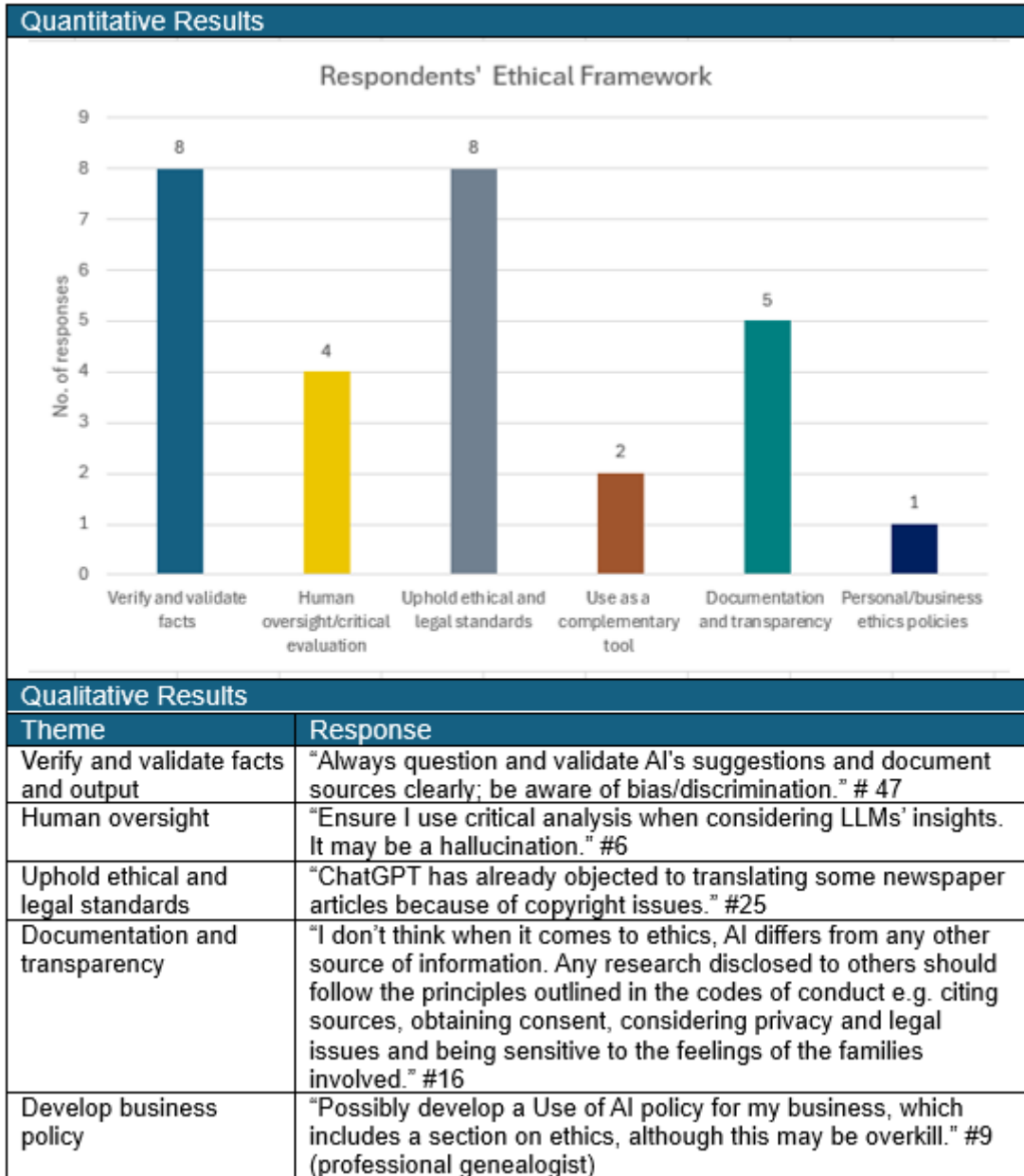


Figure 34 Ethical integration framework suggested by testing cohort.

## Continuing Education

In alignment with the literature review, the analysis indicated continuing education is crucial for addressing and mitigating ethical concerns, ensuring genealogists remain well-informed about latest technologies, ethical standards and best practices.<sup>268 269</sup> (Fig.35) These measures help prevent AI misuse, keeping research accurate, respectful, and legally compliant.

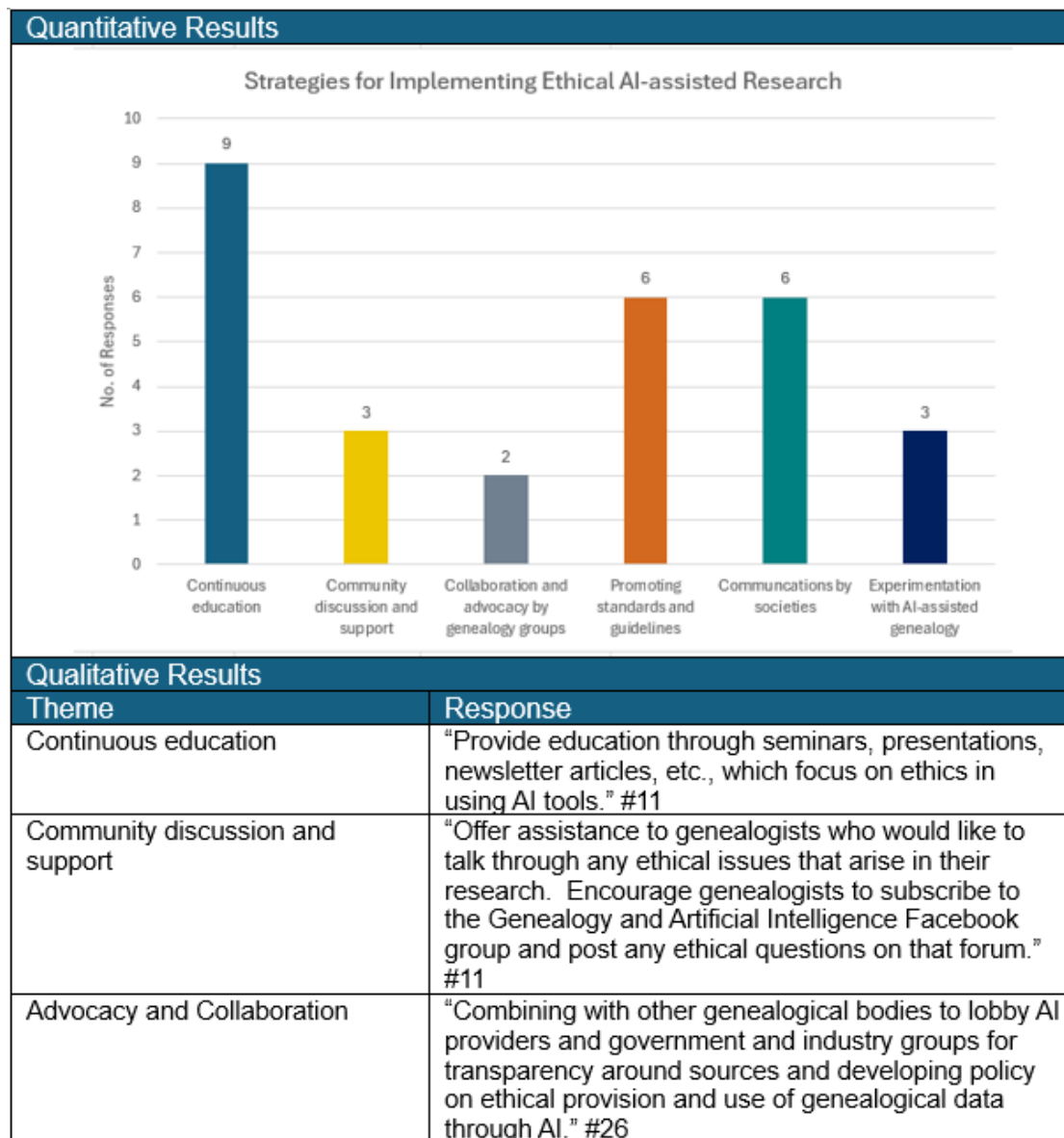


Figure 35 Respondents' recommendations for implementing ethical AI-assisted research.

Reflecting Mollick’s observation of students using AI as a sounding board, CA6 evaluated generative-AI using two ethical scenarios to assess its ability to identify potential conflicts and suggest solutions.<sup>270</sup> Results showed that AI was very effective in recognising issues and offering valuable advice. (Fig.36) Generative-AI’s ability to explore ethical problems from multiple perspectives in a non-judgmental setting is promising for genealogists and warrants further research.

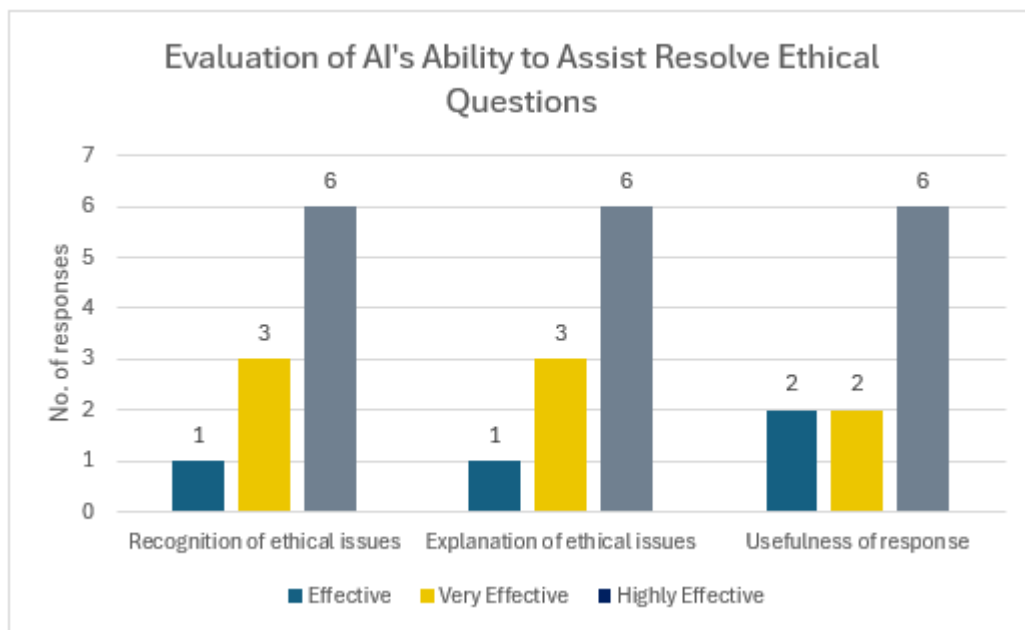


Figure 36 Evaluation of generative-AI as an assistant in resolving ethical questions.

### Risks of Inadequate Safeguards in AI-Enabled Genealogy

Neglecting ethical safeguards in AI-integrated genealogy can undermine the field’s integrity and accuracy. Existing risks, like family trees built on poor research, are compounded by the speed and ease of AI-enhanced record-matching suggestions. This often leads to the acceptance of unverified data by inexperienced genealogists. Human expertise remains indispensable for interpreting historical data. Dell’Acqua cautions against an over-reliance on technology, which can spread misinformation and inaccuracies.<sup>271 272</sup> Such oversights can misguide researchers, distort family trees, and erode trust. Additional risks include privacy violations, bias reinforcement, and intellectual property infringement, which compromise genealogical research standards.

## **Conclusion**

The analysis suggests AI can be ethically integrated into genealogical research by establishing clear guidelines for AI technologies, continuously monitoring and evaluating AI tools and fostering a culture of accuracy, privacy, and respect in genealogy communities. Mollick emphasises that today's choices about AI's alignment with human values and its potential to enhance human capabilities will have long-lasting impacts.<sup>273</sup> These issues require genealogists to take an active role in shaping the integration of AI into genealogy, a process that should begin now.

### 4.3 The Impact of AI Integration on Genealogy:

#### **RQ 4: How is AI integrated into current genetic genealogy processes, and what are the benefits and challenges?**

The literature review revealed that AI technologies, including machine learning and data analysis, are integrated into every part of genetic genealogy processes, from sample collection to the final analysis by genealogists.<sup>274</sup> The AI inherent in genetic genealogy means that by integrating DNA into traditional workflows, AI is already part of genealogy research. CA7 data on the capabilities and impacts of genetic genealogy from the end user’s perspective revealed genetic genealogy is an important complementary tool. This section blends empirical findings with analysis from literature review insights, highlighting DNA data management, AI-driven DNA analysis tools, and the broader implications for genealogy. Standard genetic genealogy terms are explained in the glossary.

#### **Prevalence of DNA Testing**

Analysis indicates that 82% of the testers have taken a DNA test. Of these, 78% have taken multiple test types or have data on multiple testing sites, showing high engagement with genetic genealogy. Reasons for not taking a test include concerns about privacy and data security, aligning with the literature.<sup>275</sup> (Fig.37).

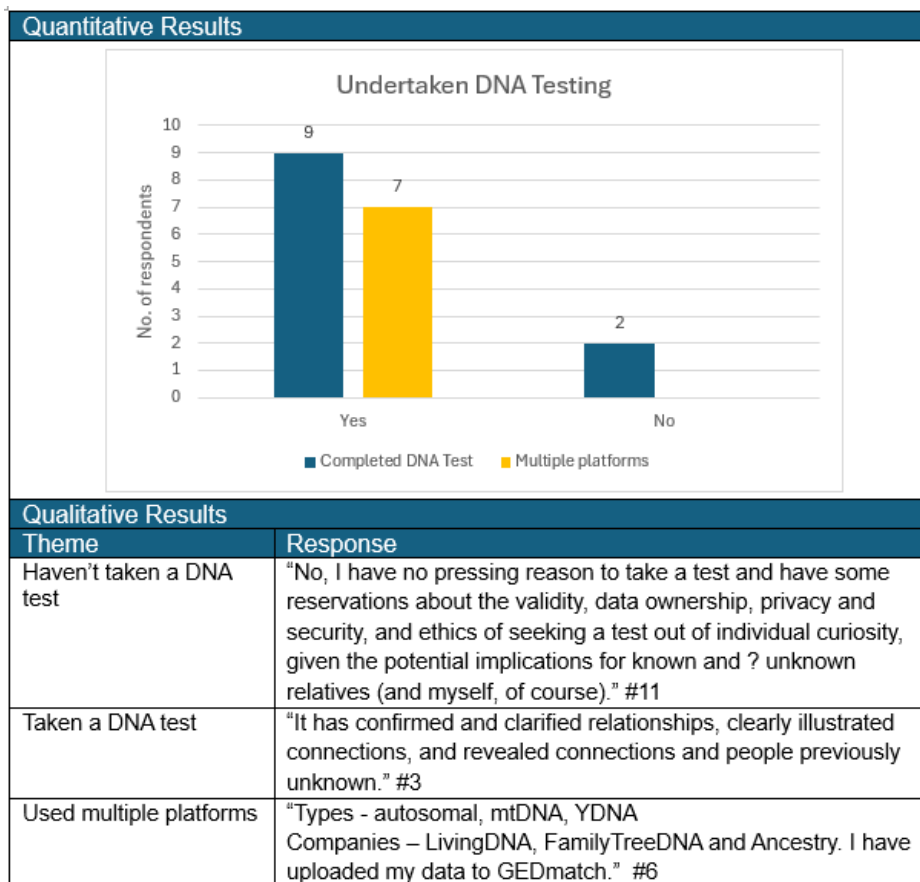


Figure 37 The prevalence of DNA testing amongst respondents.

## Integration of DNA Results into Traditional Research

The data reveals DNA testing has varied impacts on genealogy research methods. (Fig.38) While some respondents report no significant change in their research approach, others highlight DNA as a valuable tool in their research toolkit. This mixed response suggests that integrating DNA analysis into genealogy has benefits but may require more education or tools to be fully utilised by all genealogists.

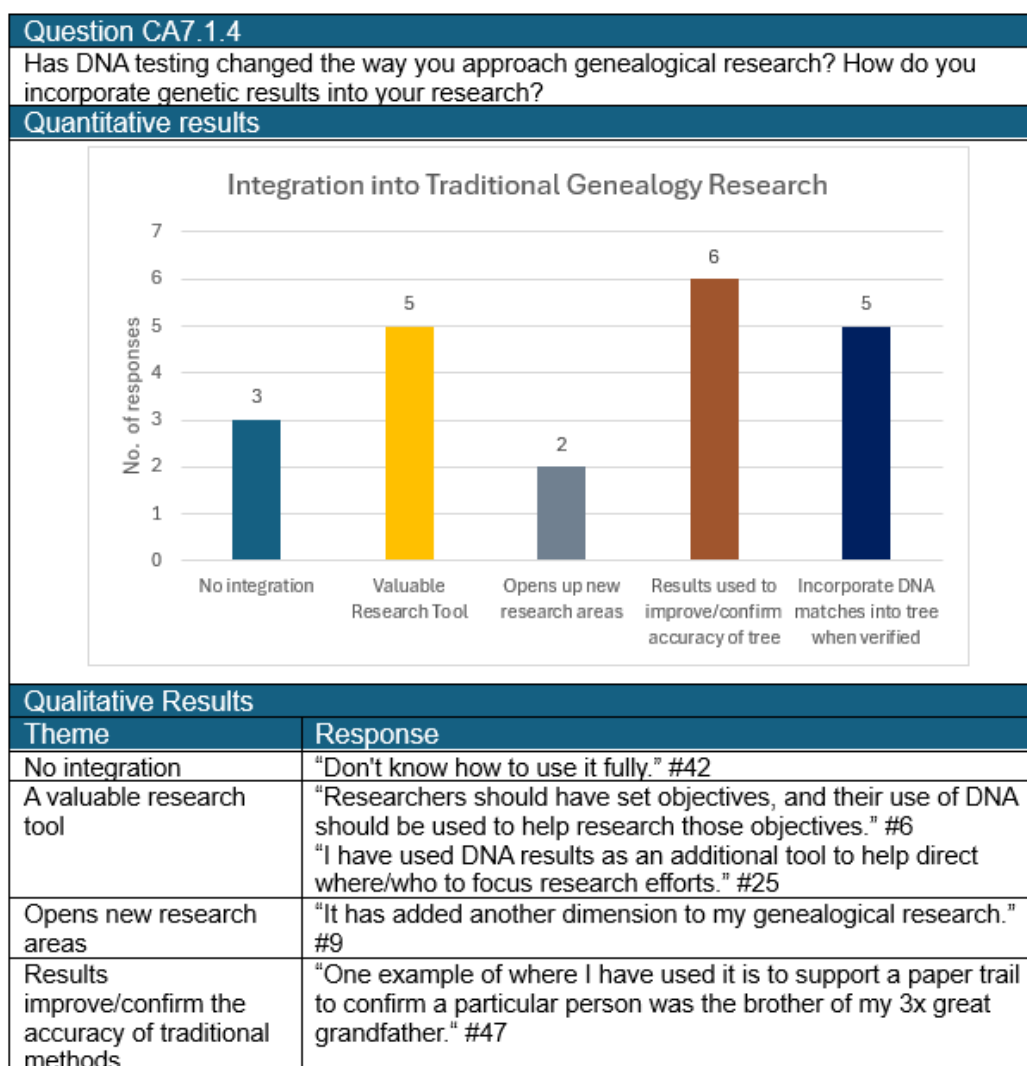


Figure 38 Results of integration of genetic and traditional genealogy methods.



The role of AI tools is central to this integration, facilitating the interpretation and incorporation of genetic data into family histories. The DNA analysis sites constantly provide new AI-enhanced tools to help genealogists analyse their results.<sup>276 277</sup> Fig.39 details the usage of DNA tools.

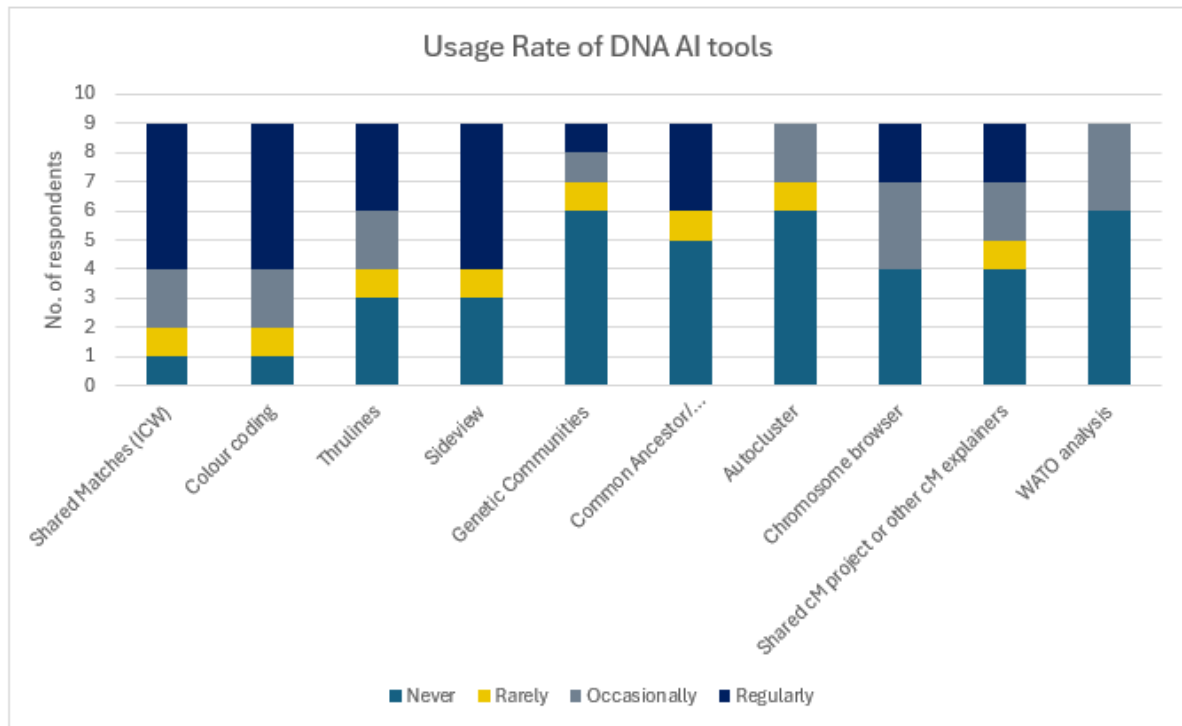


Figure 39 Frequency and usage of DNA AI tools.

The findings indicate that although there are powerful tools for interpreting genetic data, many respondents do not use these tools regularly, including fundamental ones like shared matches and cM explainers. There was a weak negative correlation between the use of DNA tools and computer skill level ( $r = -0.05$ ) and a moderate one between genealogy skills ( $r = -0.63$ ), meaning more experienced genealogists used the tools less frequently.<sup>278</sup> While this finding may indicate a trend, the small sample size limits the robustness of this analysis. An oversight during CA6 development meant respondents' genetic genealogy skill level was not collected. This would give better insights into any relationship and should be included in future research.

Several factors contribute to the underutilisation of DNA tools despite their potential:

1. Complexity and Learning Curve: Many sophisticated tools require expertise to be used effectively. The lower usage rates may result from the tools being intimidating or difficult to understand.<sup>279</sup>
2. Frequency of Necessity: Tools like Genetic Communities or Autoclusters are designed for specific analyses that are unnecessary in everyday workflows. This inherent infrequency could contribute to the lower usage rates.<sup>280 281</sup>
3. Lack of Knowledge: Unawareness of the benefits or functions of these tools means they are being underutilised. This highlights the importance of educational programs and resources to bridge this knowledge gap.
4. Perceived Complexity vs. Benefit: There may be a perception that the effort required to learn and use these tools outweighs their benefits, deterring genealogists from adopting more advanced tools.

As AI-enabled DNA analysis tools become increasingly sophisticated, it is essential to address these challenges to ensure genealogists can fully leverage their potential.

Predictions suggest the development of a DNA triangulation tool that will rapidly determine relationships between family trees within the next five years.<sup>282</sup> Although this would benefit genealogical research, poor use of the tool could exacerbate inaccurate research. Education programs and resources would help ensure genealogists can integrate these advanced technologies more effectively.

### **Benefits of AI-enhanced Genetic Genealogy**

The results indicated that 33% of the respondents reported no benefit of DNA testing for their genealogical research. The other 64% reported DNA testing confirmed existing paper trails, enabled the discovery of new family connections, and corrected previous inaccuracies in family trees. (Fig.40) These findings are consistent with the benefits of DNA testing in the literature and support Stallard's claims that the objective nature of DNA makes it a uniquely reliable source of verification.<sup>283 284 285</sup> These results emphasise AI's potential to augment traditional genealogical research with genetic information.

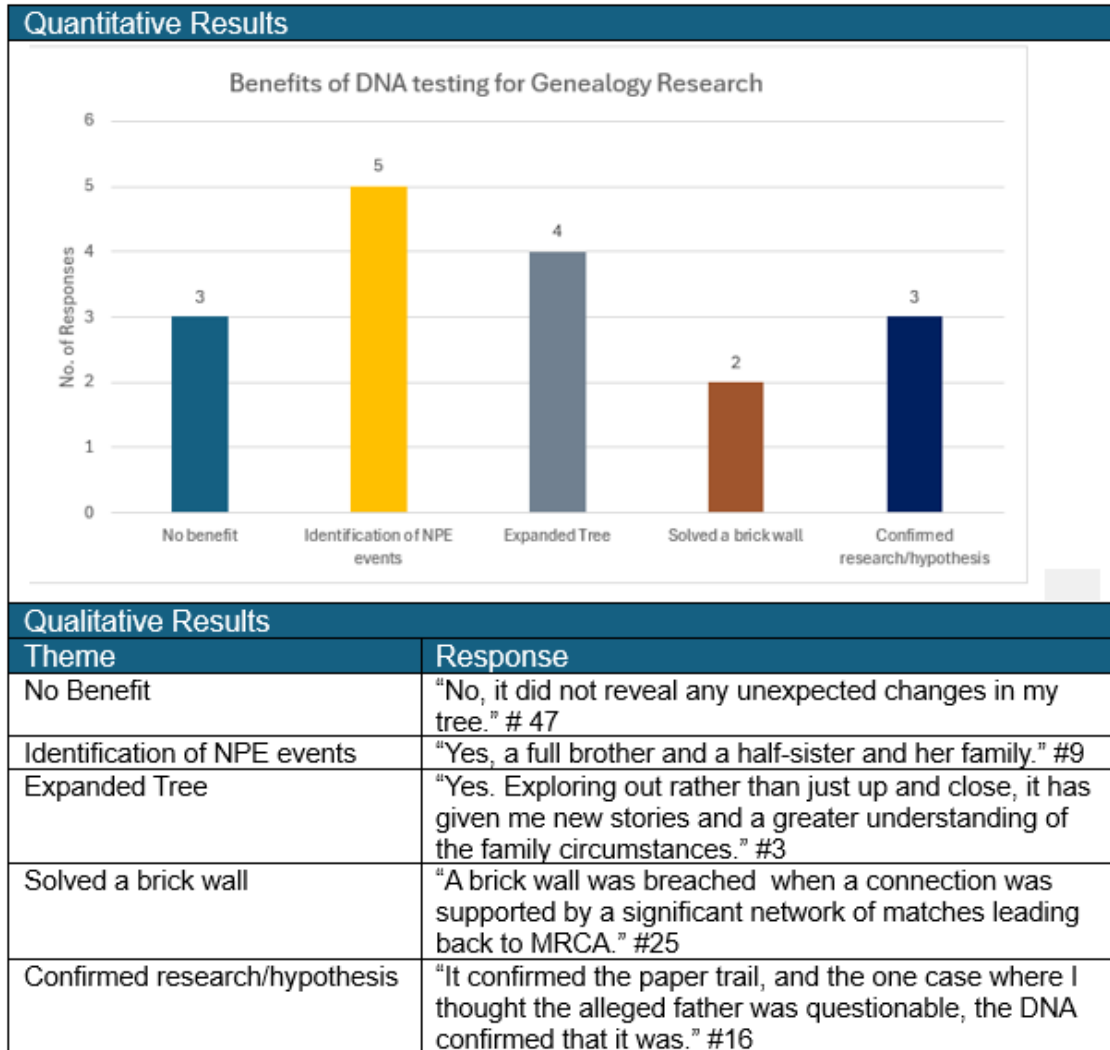


Figure 40 Results showing benefits of DNA testing.

The analysis also highlighted differences in the effectiveness and offerings of various DNA testing sites. 55.5% of respondents mentioned Ancestry as providing more matches due to its large testing pool and features, supporting Hanacek's claims that the AI algorithms and databases used by different companies can significantly impact the user experience and research outcomes.<sup>286</sup>

## Ethnicity Admixture Estimates

Although ethnicity admixture is widely advertised as the main reason for taking a DNA test, 55.4% agreed with Holton and Stankus' assertions that its use for genealogy research is of limited benefit and results must be treated cautiously.<sup>287 288</sup> Interestingly, the four respondents who found it valuable had non-British ethnicities in their admixture. (Table.13)

Qualitative and Quantitative Results		
CA7.1.6 How have you used your ethnicity admixture estimate results in your research?		
Theme	No. of Responses	Qualitative Feedback
Haven't used	5	<p>"There is nothing in it that I didn't know without a DNA test". #22</p> <p>"The results are only estimates. I get different admixtures at Ancestry, Family Tree DNA and Living DNA which are different from my ethnicity result from document research. The companies update from time to time, and interestingly, they are getting closer to my research results. So, I use my research results." #6</p>
Used	4	<p>"To confirm Portuguese connection, which was done via an uncle, which appeared/disappeared when Ancestry updated estimates." #26</p> <p>"To support family folklore that one of my Irish ancestors fought with the French against the English in the late 1700s#26.d married a French woman before returning to Ireland." #9</p>

Table 13 Use of ethnic admixture results

## Challenges of AI-enhanced Genetic Genealogy

The main concerns raised regarding AI-enabled genetic genealogy are consistent with issues of privacy, ethical use, and discrimination documented in the literature.<sup>289 290 291</sup>

(Fig.41) The respondents' acknowledgement that AI tools should be combined with human expertise and validated outcomes for accuracy is consistent with the principles outlined in the Genetic Genealogy Standards and the ISOGG Code of Practice.<sup>292 293</sup>

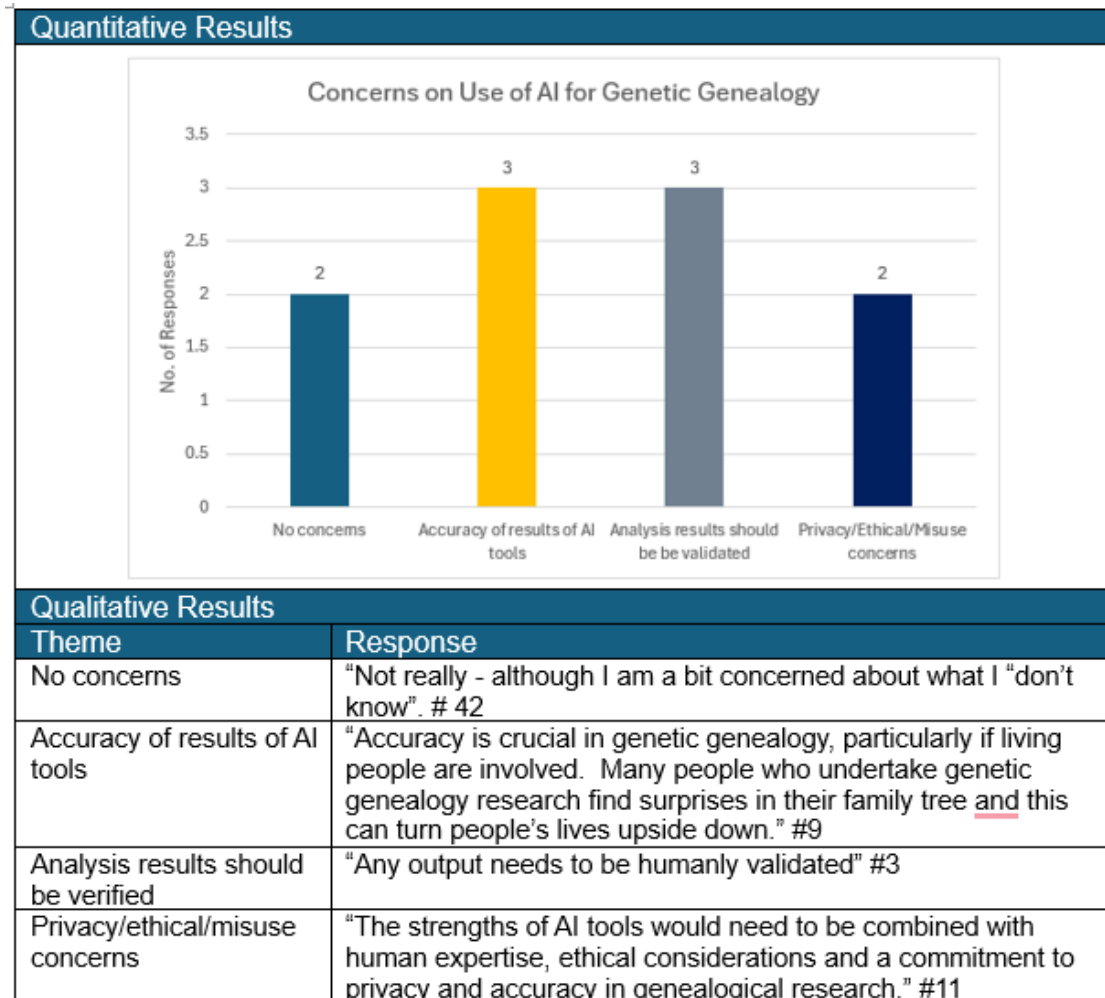


Figure 41 Evaluation of concerns about the use of AI in genetic genealogy.

## Conclusion

These results indicate that AI already plays a significant role in genetic genealogy, with the potential for greater impact as technology evolves. Key benefits include increased efficiency and accuracy in data analysis and interpretation. At the same time, challenges involve ethical use, data privacy, and the ongoing need for human expertise to guide and contextualise AI's insights. The results also emphasise the importance of continuing education and tool development to maximise the benefits of genetic genealogy.

### ***RQ 5: How is AI changing the availability, access and use of historical documents?***

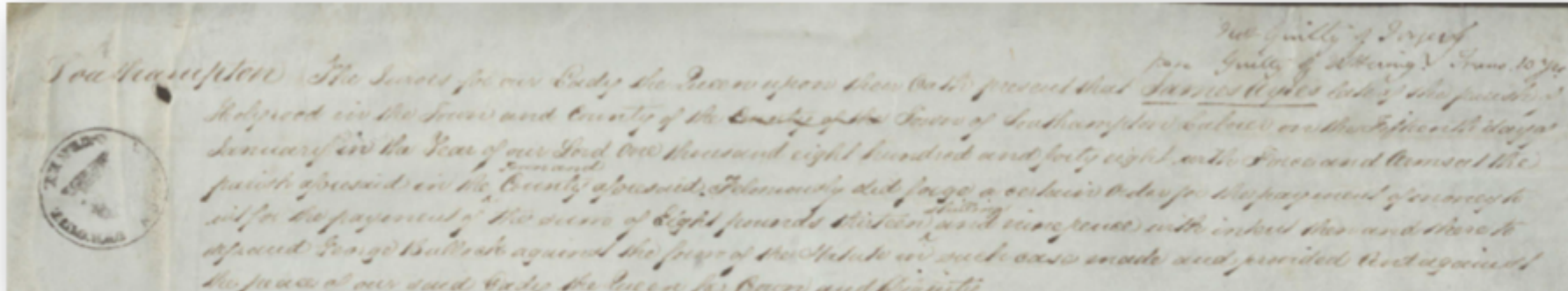
The literature review showed integrating AI technologies in libraries, archives, and genealogical platforms enhances genealogists' ability to rapidly search large datasets and connect information across multiple records.<sup>294 295</sup> CA8 collected data on the impact of this technological change on genealogists' research methods. This section investigates how AI technology empowers genealogical research by blending CA8 results, design-phase evidence and literature review insights.

#### **Increasing Historical Document Availability**

Genealogy websites use AI to digitise billions of records annually, consistently expanding documents accessible to genealogists.<sup>296</sup> Now, AI handwriting text recognition algorithms (HTR) are accelerating the availability of these records for genealogical research.<sup>297</sup>

Although high-quality HTR systems are costly, complex and mainly limited to large organisations, consumer-accessible AI tools can be used for handwriting transcription.<sup>298</sup> For instance, the project's design-phase focus group demonstrated transcribing a three-page handwritten court indictment using common AI tools in under seven minutes, a task estimated to manually take at least a day. (Fig.42) Although this process is imperfect and required experience in using AI, subscriptions to Transkribus and ChatGPT4, and human curation to produce valuable results, it highlights AI's potential for this task. Consumer-accessible HRT systems are rapidly improving, and experts predict that generative-AI will significantly enhance HRT capabilities within the next five years.<sup>299</sup>

Original enhanced with MyHeritage Photo Enhancer



**Enhanced image to Transkribus**

- Not Guilty of Jorjery
- pose Guitty of Lettering Truns 10 yr.
- You hampton The Jurors for our Lady the Queen upon their cash present that sames Uyles lase of the parish Holirood in the Town and county of the Cieity of the Town of Southampton Caloieron the Fifteenth day o

**Transkribus to ChatGPT 4.5**

Not Guilty of Forgery  
Plea: Guilty of Uttering. Sentenced to 10 years

You hampton, the jurors for our Lady the Queen, upon their oath present that James Ayres, late of the parish of Holyrood in the town and county of the city of Southampton, laborer, on the fifteenth day of January in the year of our Lord one thousand eight hundred and forty-eight, with force and arms at the parish aforesaid in the county aforesaid, feloniously did forge a certain order for the payment of money,

Figure 42 The stages of converting handwriting to text using consumer-level AI.

## **Searching Records in New Ways**

As more digitised resources become available, the limitations of traditional searching methods become apparent.<sup>300</sup> Just as computer indexing replaced card catalogues, AI-enabled full-text searching and Natural Language (NL) searching will explore entire documents rather than prescribed fields.<sup>301</sup> This shift empowers genealogists to uncover connections and insights previously hidden in unindexed records.

Currently, NL searching of genealogy records is only available at MyHeritage, while full-text searching is in beta mode at FamilySearch Labs.<sup>302 303</sup> Only NL searching was tested in CA8, as full-text searching became available in March 2024 after data collection was finished.

Analysis of CA8 findings showed while respondents found NL searching using a chatbot intuitive and relevant, qualitative comments revealed some dissatisfaction with the app. (Figs.43-44) This discrepancy highlights the importance of considering both quantitative and qualitative data for a comprehensive understanding of user experiences.<sup>304</sup>



# MyHeritage AI Record Finder<sup>®</sup>

To initiate your genealogy search, kindly provide the first and last names of the individual you are researching, along with any pertinent details such as life events or close relatives' names.

Hey there, search buddy! 🗨️ I'm your genealogy sidekick. Just give me a name and any extra deets (dates, places, relatives) and let's dig up some family treasure! 🥳 🌳 📖

Above: Users can choose between a formal and informal style interface with the chatbot  
Below: Example of the limited results available.

The screenshot displays a chat interface. At the top, a green bubble contains the text: "John Drake born in Yorkshire around 1862. Married Mary Chapman 11 feb 1889". Below this, a grey bubble says: "Here are the results found for the information you provided:". Three result cards are shown below. The first two are orange and feature a family tree icon, with the text "John Henry Drake" and "MyHeritage Family Trees". The third is a dark card with a census document image, titled "John Drake" and "1881 England & Wales Census".

Figure 43 Example of a MyHeritage Natural Language search.<sup>305</sup>

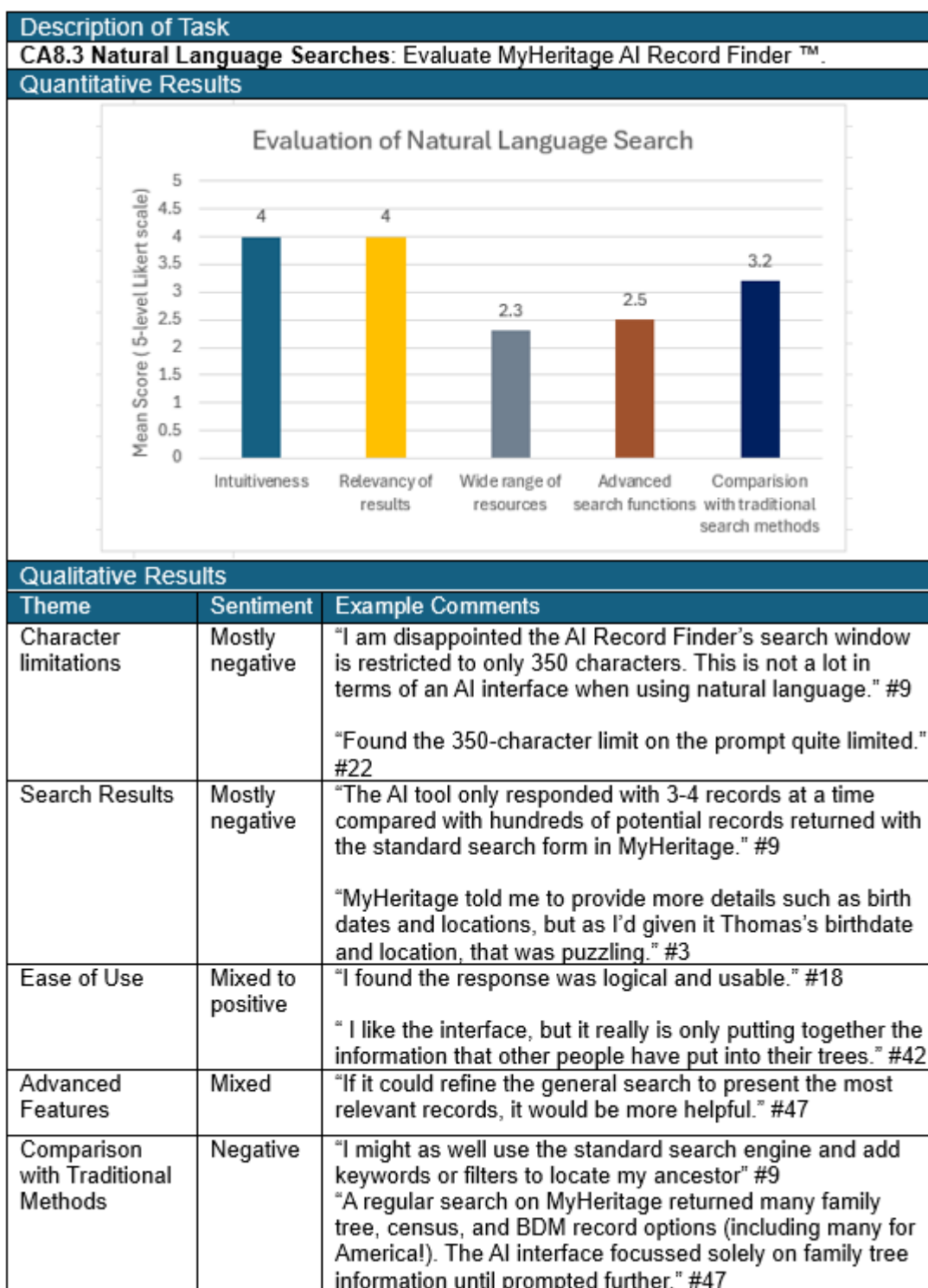


Figure 44 Evaluation of MyHeritage AI Record Finder.

Respondents rated features like intuitiveness and relevancy highly based on app functionality. However, scores for essential features like searching a wide range of records and advanced search filtering were low, and the chatbot did not offer any benefits compared to traditional search interfaces. Qualitative feedback allows respondents to provide valuable insights into areas needing improvement and express their thoughts in detail, including

frustrations or specific issues, like the small character window. <sup>306</sup>This detailed feedback can lead to more nuanced observations than quantitative ratings alone.

Many of the issues users identified are most likely caused by NL searching being an emerging technology. As it matures, the character window should increase, enhancing search capabilities. The app's performance was influenced by the chatbot's processing power and the limited datasets in the beta version, explaining the limited record returns. Improvements in these areas should lead to better performance and more accurate results. User feedback helps tune the training data of NLP systems, refining and optimising AI's performance.<sup>307</sup> The search comprehensiveness and accuracy will improve as the system expands to include more record collections and learns from user-feedback.

FamilySearch full-text searching is also currently limited to a few record collections. Facebook feedback is generally favourable; hopefully, this interface will be rolled out to more record collections.<sup>308</sup> Future research could reevaluate these findings after the technology is more widespread, offering new insights into these new AI-enabled search methods.

### **Locating and Accessing records**

CA8 gathered data on respondents' strategies for locating a UK birth record based on information in related documents. Analysis revealed all respondents started with online research, a significant shift in genealogical research methods, which aligns with Lucy's research.<sup>309</sup> Online databases' convenience and accessibility facilitate this change. AI technologies embedded in genealogy websites enhance this experience with sophisticated record linkage and filtering capabilities, expediting the research process and revealing connections otherwise difficult to find.<sup>310</sup> Features genealogists take for granted, like record hints and related records, rely on machine learning and neural networks, which analyse large amounts of data to find potential matches, suggest records and identify patterns that might not be immediately obvious to human researchers.<sup>311</sup> Genealogists can improve these features by providing feedback on inapplicable record suggestions, helping systems to learn and select more relevant documents.

Analysis revealed search strategies varied by genealogical skill, with the beginner struggling with the task, while expert genealogists used a broader range of records, including non-digitised resources. This suggests expert genealogists better understand digital databases' limitations and the value of diverse sources. (Fig.45).

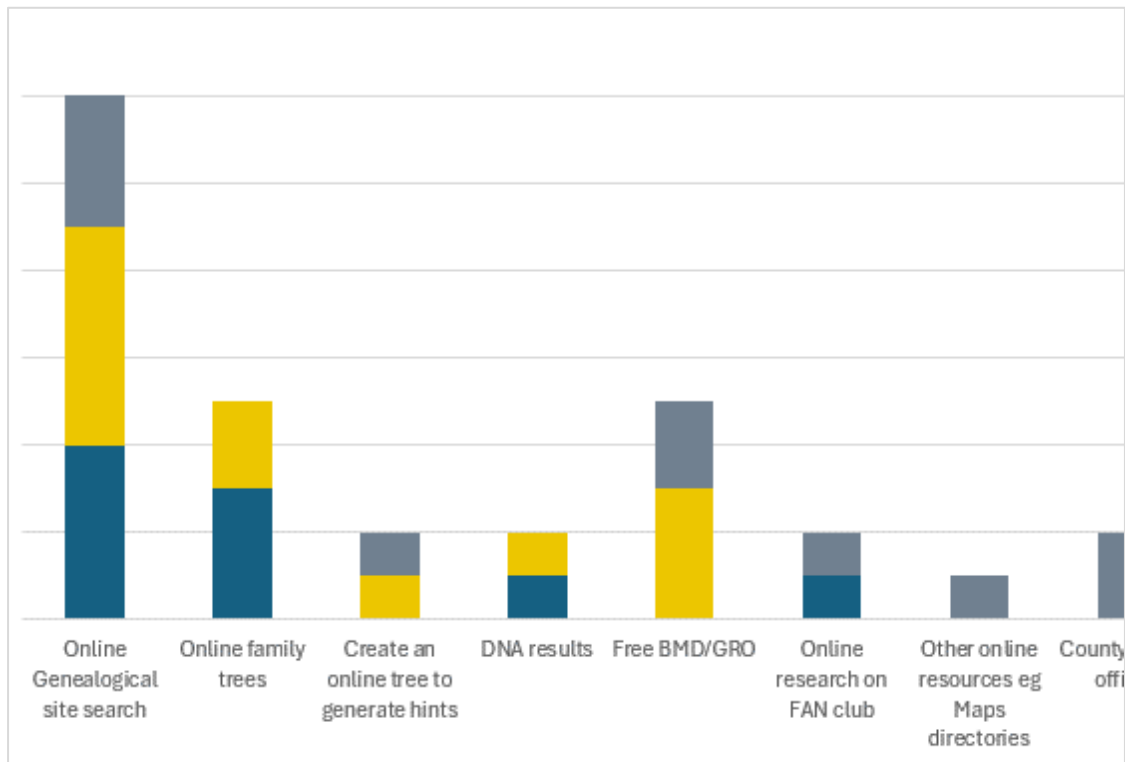


Figure 45 Search strategies used by respondents to find a UK birth record.

Generative-AI can help minimise these differences by creating research plans or strategies to learn new research skills. (Fig.46) These findings align with Duff's conclusions on genealogical research's depth and complexity.<sup>312 313</sup>

## AI-generated Research Plan

### **1. Estimate Birth Year and Location:**

John Drake was 25 at the time of his marriage in 1889, which suggests he was born around 1864. Given his marriage took place in Warrington, Lancashire, it's possible he was born in or near this location, though his profession and his father's could imply they moved around.

### **2. Search Civil Registration Records:**

Start by searching the civil birth registrations in England and Wales from 1863 to 1865 to account for any inaccuracies in age reporting at the marriage. Use websites like FreeBMD, GRO Online, or Ancestry.co.uk, focusing initially on Lancashire and then expanding geographically if needed.

### **3. Census Records:**

Examine the census records for 1871, 1881, and possibly 1891 to locate John Drake and gather more information about his family structure, movements, and potential siblings. This might help confirm his father's name and profession, ensuring you have the right family, especially given the alterations noted in the marriage document.

### **4. Parish Registers:**

Since civil registration might not capture everyone, especially in earlier years, check the parish registers for baptisms around 1864 in Lancashire. Given the profession of earthenware dealer, it's possible the family was not highly mobile initially. Local archives or online church record databases like those on FindMyPast could be useful.

### **5. Probate Records:**

Since there was a note about George Drake being deceased but then crossed out, it might be useful to check probate records around the estimated time of his death. This might provide clues about family locations or confirm identities.

### **6. Address Consistency:**

Check directories and electoral rolls for any listings at 33 Buttermarket St in Warrington around the time of John's marriage

### **7. Professional Associations:**

Since both John and his father were earthenware dealers, there might be trade directories or guild records

### **8. Broadening the Search:**

If searches in Lancashire are inconclusive, consider the possibility that the family might have moved from another part of the UK, given the mobility suggested by his son's birth in Sheffield and the profession listed. Regions with known pottery industries like Staffordshire could be potential areas of origin.

Figure 46 Example of an AI-generated research plan to locate a UK birth record (CA8.2) <sup>314</sup>

Respondents rated the relevance of AI-created plans highly, with an average score of 4.2/5. The depth of suggested records received a strong score of 4.4/5, indicating substantial thoroughness. (Fig.47)

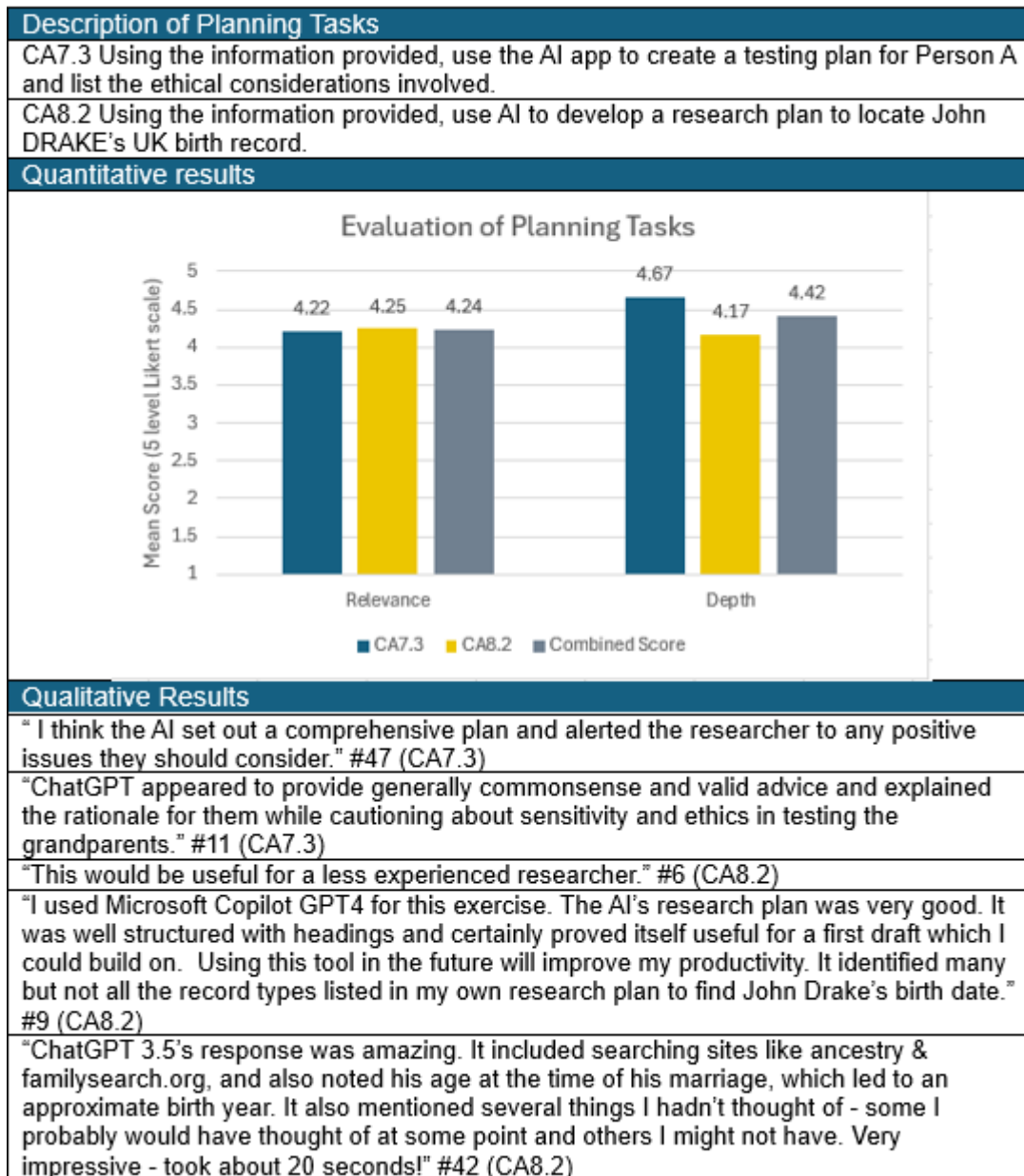


Figure 47 Evaluation of AI-generated research plans.

The findings support Lucy’s observations on technology’s significant role in transforming genealogical research.<sup>315</sup> Both studies noted a shift from traditional research methods involving physical archive visits to digital and AI-enhanced research, allowing for more efficient and expansive research, particularly for non-local records. The findings also support Lucy’s comment that genealogists will use any method or technology available to access a record.<sup>316</sup> (Table.14)

What would you do if the record was not available on line? n=12						
Action	1st Choice	2nd Choice	3rd Choice	4th Choice	5th Choice	6th Choice
Alternative on-line documents	9	3				
Books, fiche/film, manuscripts	1	7	2	1		
Contact Archive			6	3	2	
Visit Archive	1	1	1	6	3	
Postpone search		1	2	1	6	2
Other	1				1	2

Table 14 Heatmap showing respondents’ options if a record was unavailable online.

### Limitations - Biases

While the shift towards digital methodologies offers significant advantages, it has limitations. The literature review suggests digital collections often have biases towards mainstream research interests.<sup>317 318 319</sup> Qualitative analysis revealed respondents' perceived bias in online genealogical databases, particularly regarding record availability and source and potential biases by larger genealogical companies towards English-speaking regions. (Fig.48) These insights highlight the importance of Milligan's advocacy for better education on digital search methodologies to help users understand the subtle biases shaping digital archives.<sup>320</sup>

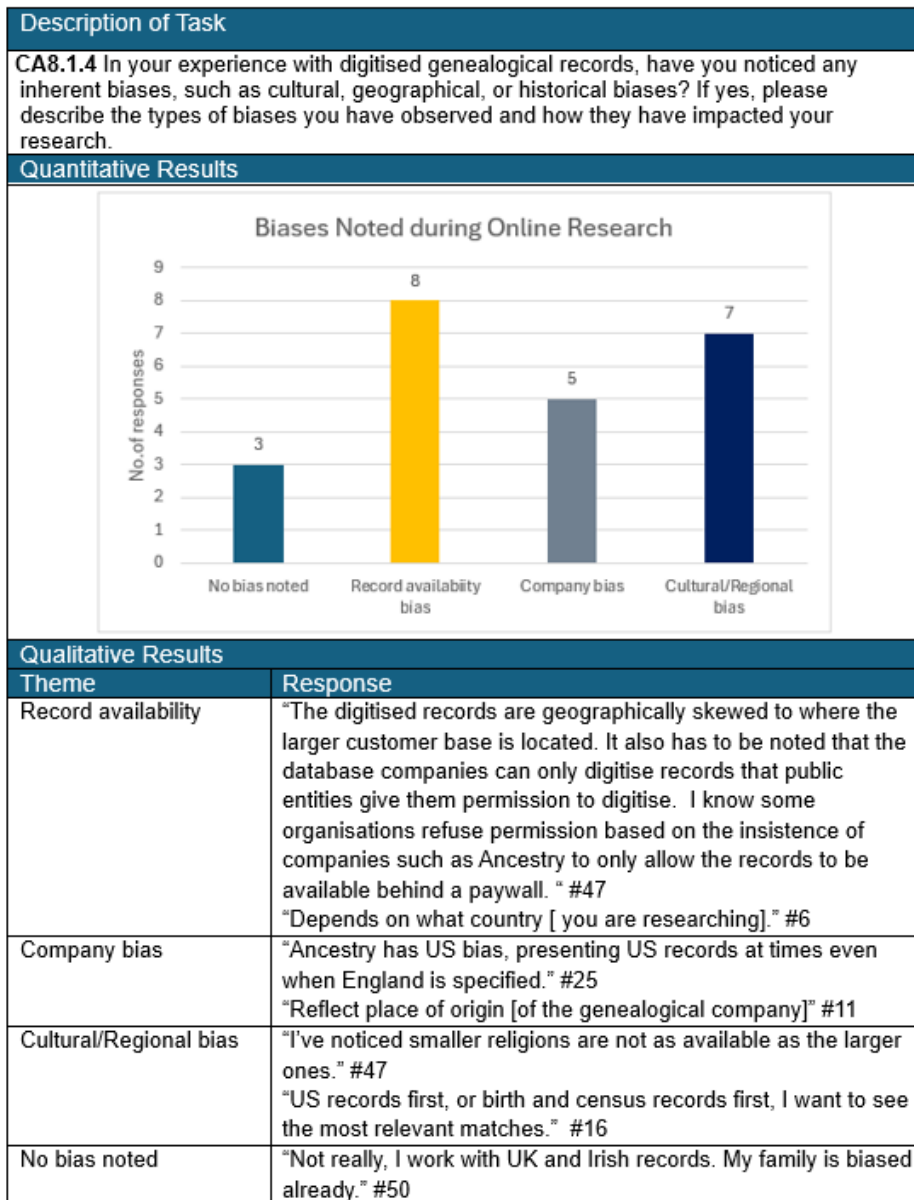


Figure 48 Evaluation of bias in online research.

## Conclusion

The analysis indicates AI technologies are impacting genealogists' research methods, empowering them with better access to historical documents than traditional methods. The findings support the literature insights emphasising the effectiveness of AI technology in making new material available, improving document accessibility and making it easier to find elusive records. This impact will continue as AI technologies develop to handle the ever-increasing volume of digitised documents, stressing the importance of ongoing education in working with digital methodologies.



**RQ6: How is AI being used to re-interpret historical data to provide new insights for genealogical research?**

A different research approach was used to investigate three research projects using AI-enhanced systems to reinterpret convict records. AI technologies in each project enabled new interpretations of historical data and innovative methods to share the results. Table.15 provides a brief overview of the projects with detailed information in Appendices.F and G. Data was collected through informal interviews with project leaders focusing on the historical impact, AI usage and AI-related challenges. Additional insights came from academic papers and workshop presentations by the project leaders.<sup>321 322</sup> This section presents the empirical findings, integrating descriptions and analysis with literature review insights.

	Project 1 Ironclad Sisterhood	Project 2 Landscape of Production and Punishment	Project 3 Making Crime Pay
Project Overview	A small-scale project driven by genealogical society members to re-interpret the lives of convict women in New South Wales 1788-1818.	A large-scale multi-disciplinary academic project mapping the changes in convict management on the Tasman Peninsula (Tasmania).	A large-scale multi-disciplinary academic project researching the connections between convicts and Tasmanian colonial society.
Historical Impact	New insights into the contributions of convict women to the development of colonial Australia. New insight into women's lives at a time when most records relate to men. Challenged stereotypical views of convict women as prostitutes.	Mapped changes in the landscape against convict management changes. New insights into the work, management and punishment of convicts.	New insights into the daily lives of convicts. New insights into colonial/convict networks. Creation of <i>Unshackled</i> Convict Memorial, immersive experience. Challenged stereotypical views of convict life.
Types of AI used	Database analytics.  Digital storytelling. Image/avatar creation.	Advanced database analytics. Spatial analysis – GIS, LIDAR.	Advanced database analytics. Spatial analysis- GIS, LIDAR. Digital Storytelling. Image/avatar creation.
Level of AI used:	Consumer-level AI	Bespoke or commercial-level AI	Bespoke or commercial level AI
AI Challenges:	Inbuilt bias in image generation distorting accuracy. Need to attribute images as AI-generated.	Mitigation of record selection bias. Ensuring AI technologies were integrated into traditional archeological and historical research methods.	Mitigation of record selection bias. Difficulty with the creation of realistic avatars. Mitigation of incomplete datasets

Table 15 Overview of AI-enhanced projects to reinterpret historical convict records.

## **Project 1: *Ironclad Sisterhood***

### Project Overview

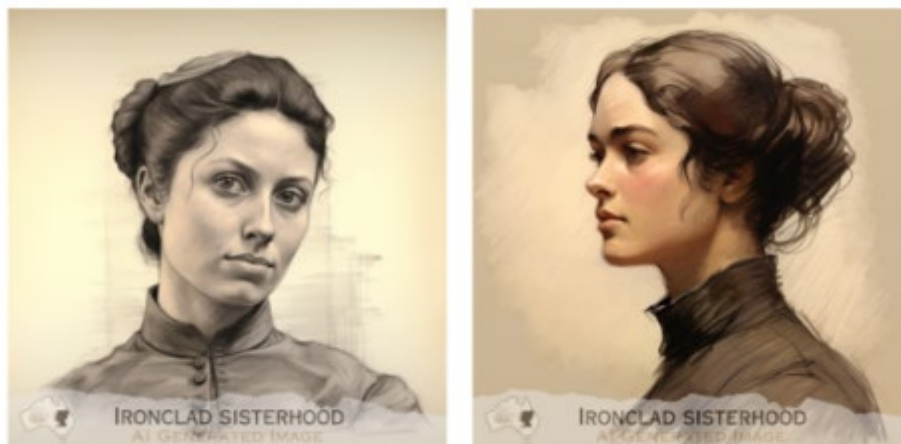
Led by Redfern, this project used consumer-level AI and digital tools to re-examine an overlooked historical dataset, offering fresh perspectives into the experience of female convicts.

### Use of AI technologies

It employed AI-assisted database management, text-to-image avatars, and AI-generated content creation for biographies and podcast scripts, democratising access for a broad audience as advocated by Sharratt and Crymble.<sup>323 324</sup> This approach aligns with digital history scholars' support for integrating technology into historical research.<sup>325 326</sup>

### Challenges and Ethical Concerns

The project encountered challenges with text-to-image generation, specifically, a tendency of AI to “beautify” subjects by omitting scars and other blemishes described in the convict database, potentially distorting historical accuracy. (Fig.49) This may be due to Generative-AI’s guardrails preventing scars and disfigurements on avatars due to bias in training data favouring unblemished faces or design choices aiming for visual appeal and user comfort.<sup>327</sup>



*Figure 49 "Beautification" of convict and acknowledgement of AI-generated images.<sup>328</sup>*

This issue, alongside the ethical concern of disclosing AI involvement, underscores the delicate balance needed between AI’s creative potential and its limitations in representing historical truths accurately. Redfern addressed the issue by adopting the Genealogy and AI Facebook group guidelines and acknowledging the use of generative AI on the images.<sup>329 330</sup> These concerns resonate with Crymble's call for ethical mindfulness in digital history, highlighting the need for stringent ethical guidelines to navigate the challenges of bias and accuracy in AI-enhanced historical research.<sup>331</sup>

## Tasmanian Longitudinal Dataset (TLD)

The TLD, an AI-enhanced digital archive encompassing over 1.6 million records from the 19th and early 20th centuries, is central to the following two projects. The use of its curated contextual datasets for historical and genealogical research aligns with the work of Godfrey, Gibbs, and Milligan, who highlight digital archives as essential resources.<sup>332 333 334</sup> By digitising and linking records such as convict records, prison and hospital admissions and military enlistments from multiple sources, the TLD preserves the records' original context, supporting Cowley's research.<sup>335</sup> (Table.16)

Category	Data Source	Details
Convict Records	Conduct records, convict death registers	Information on 64,819 male and 13,673 female transported convicts, including detailed punishments
Colonial Courts and Criminal Justice	Tasmanian Supreme Court cases (1824–1939), lower court records, Tasmanian Police Gazette, Prosecution Project Database	Records of defendants, charges, and verdicts, notices of prisoners
Census and Musters	Annual convict musters, household censuses (1837–1857)	Data on household occupants and properties, including census returns
Departures and Arrivals	Free arrivals and departures records, assisted migrant records	Details of 24,232 arrivals and 114,452 departures, records for assisted migrants (1852–1858)
Births, Deaths, and Marriages	Civil registration records, ecclesiastical records	195,000 births (1838–1899), 51,000 marriages, 155,000 deaths (1838–1928), baptisms, and burials
Property and Financial Records	Convict Savings Bank records, Hobart Savings Bank records, trade and street directories, property valuation rolls	Financial records, addresses, and property valuations for Hobart and Launceston
Hospital and Pauper Admissions	Hobart Hospital deaths, St Mary's Hospital admissions, New Norfolk Psychiatric Hospital admissions, pauper admissions and discharges	Detailed hospital and pauper admissions and discharges
Military Records	WWI attestation papers	Data on 15,234 Tasmanian-born soldiers and nurses, including personal details
Maps and Plans	Tasmanian Archives, Land Tasmania	Exploration charts, architectural plans, title deeds, and other cartographic materials
Meteorological Data	Daily weather measurements for Hobart	Minimum and maximum temperatures, barometer readings, and wind direction since 1825

Table 16 Example of TLD datasets.<sup>336</sup>

This diversity of data helps mitigate the data selection bias raised by Little, resulting in a more accurate representation of Tasmania's colonial history.<sup>337 338</sup> The TLD was upgraded with cutting-edge AI database management technologies as part of Project 3, and ongoing funding from the project was allocated to support its maintenance.<sup>339 340</sup>

## Project 2: Landscape of Production and Punishment

### Project Overview

Tuffin's project focused on the innovative application of AI-enhanced LiDAR and Global Information Systems (GIS) technologies to uncover the impact of changes in convict management on convicts' lives and the Tasmanian landscape.

### Use of AI technologies

LiDAR and GIS helped create 3D archaeological site models, providing new insights into convicts' historical land use. (Fig.50)



Figure 50 LiDAR-generated elevation map of Port Arthur showing significant activities.<sup>341</sup>

## Findings and Insights

The LiDAR maps revealed unseen features that provided insights into the convict's land use. For example LiDAR imagery at the Convict Quarry revealed several tramways, not visible at ground level, evidence of sophisticated land management and transport systems. (Figs.51-52).



Figure 51 The Port Arthur Convict Coal Mines showing terrain visible from ground level <sup>342</sup>

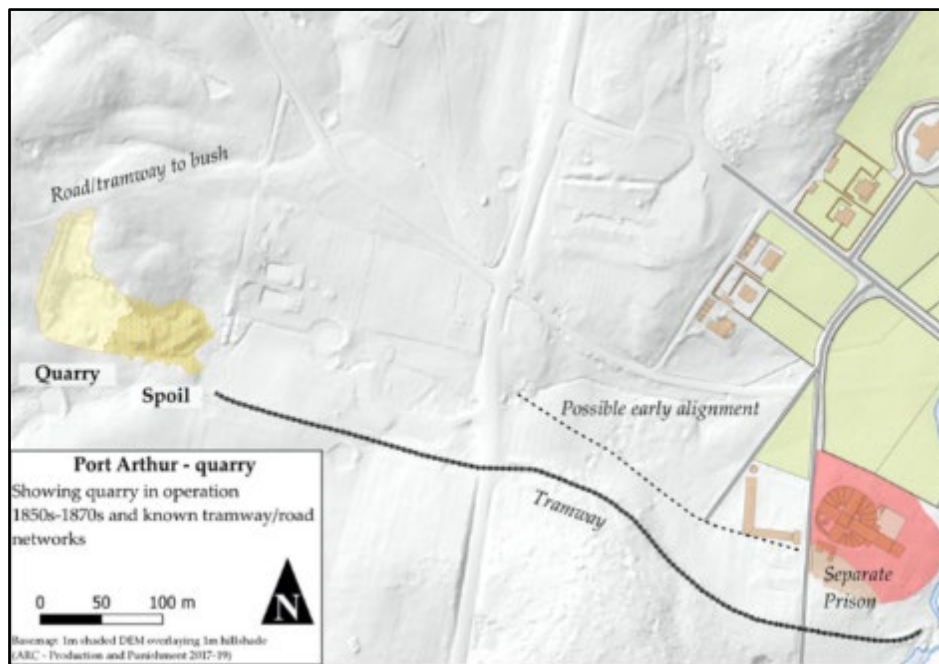


Figure 52 LiDAR scan of the quarry site revealed a spoil heap and tramlines. <sup>343</sup>

Convict punishment data from the TLD was mapped onto these LiDAR maps using GIS technologies. This mapping allows researchers to see where crimes occurred, the number and type of crimes at that location and link them to the source documents. Fig.53 illustrates the link between a convict's conduct record, the place at Port Arthur where it occurred and the other crimes at that location. This information adds valuable context when researching convict ancestors and helps contextualise the records to provide new insights into their lives.

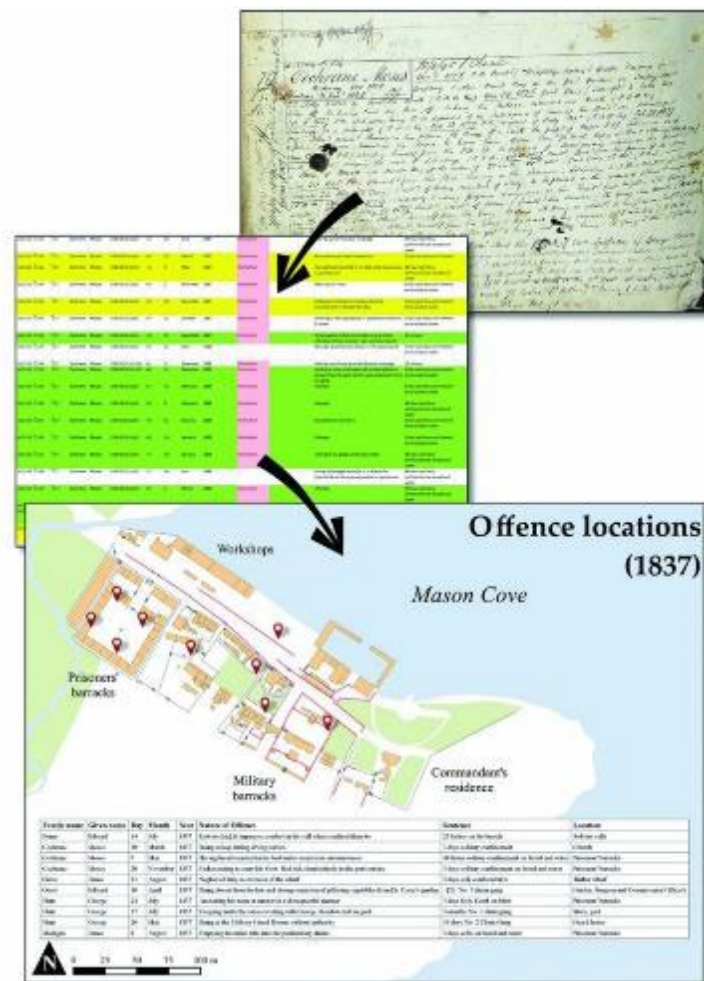


Figure 53 The process from convict conduct record to table and GIS geo-located map.<sup>344</sup>

This project embraces the methodological innovations discussed by Milligan and showcases how digital technologies can deepen our understanding of historical contexts.<sup>345</sup> The project balanced technology with historical accuracy, a fundamental element of Milligan's research, by employing a multidisciplinary approach mirroring the recommendations of Crymble, resulting in minimal AI-related challenges.<sup>346 347</sup>

### **Project 3: Making Crime Pay: Manufacturing Heritage Experiences in a Digital Age**

#### Project Overview

Maxwell-Stewart's large-budget project used the TLD and advanced AI technologies to reinterpret Tasmania's convict heritage. This multidisciplinary approach created immersive digital heritage experiences and enhanced visitor interactions, significantly conserving and interpreting Tasmania's convict history.

#### Interactive Memorial

The project created an interactive convict memorial hub at the Hobart Penitentiary Chapel (HPC), Tasmania's most significant convict site and administrative centre for convict labour from 1821.<sup>348</sup> Launched in November 2023, the interactive memorial "Unshackled" educates the public on convicts' roles in Tasmania's colonisation and offers a deeper understanding of their lives.<sup>349</sup> Visitors use smartphones to search the TLD and adopt a convict whose details, including an AI-generated avatar, crime, sentence, voyage, and life in the colony, appear on four large-screen displays suspended from the Chapel roof. The project also updated the TLD with the latest AI database technologies, ensuring its sustainability with ongoing funding from the exhibition. (Figs.54-55)



*Figure 54 The Convict Memorial "Unshackled" at Hobart Penitentiary Chapel.*

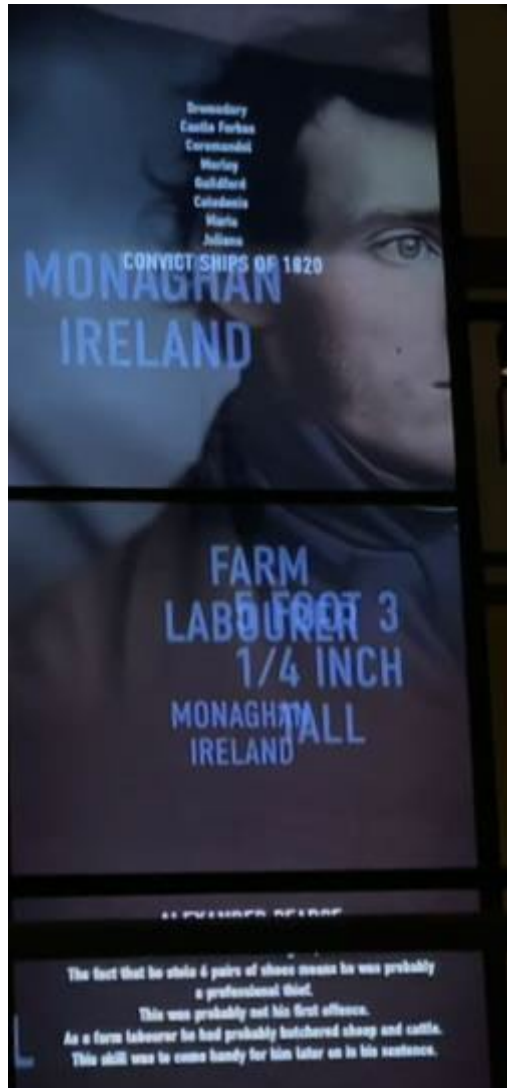


Figure 55 Close-up of one side of the memorial pillar in the *Unshackled* immersive experience.<sup>350</sup>

Plans are underway to expand the project to other Tasmanian convict sites, offering location-relevant information. For example, at the Ross convict bridge, the app provides additional contextually relevant details on the adopted convict.

### Use of AI Technologies

Maxwell-Stewart's team used AI applications to create "Unshackled," incorporating innovative technologies to bring convict histories to life. Spatial analysis using LiDAR and GIS mapped and analysed convict sites, and AI-generated convict avatars created realistic representations based on TLD descriptions and existing images of convicts. (Fig 56) Future enhancements will include pictures of convict descendants in the training data to improve the image's realism.



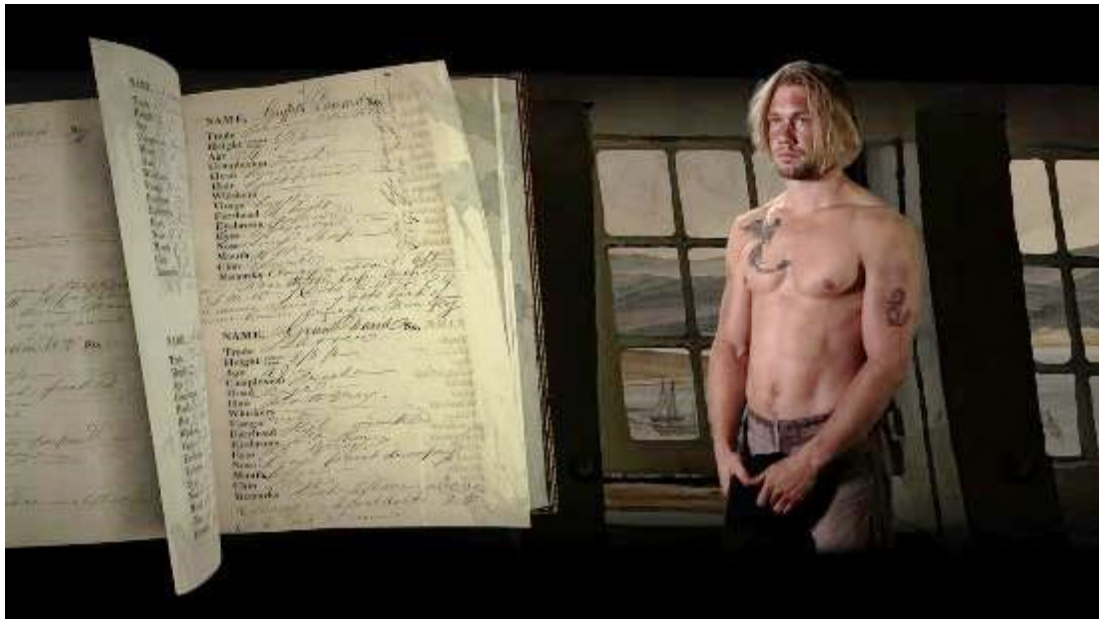


Figure 56 Convict record of Edward GRIFFITHS and digital avatar made from description. Note the inclusion of tattoos. <sup>351</sup>

The project also utilised thematic digital storytelling, employing computer scripts to visualise the lives of 78,000 convicts by identifying event sequences, locations, and interconnected storylines through network analysis. Scheherazade-IF, an AI approach, was used to fill narrative gaps in minimally documented convict stories with plausible thematic content. <sup>352</sup> This exhibition narrative can also be adapted for different audiences, including school groups.

### Challenges and Ethical Concerns

The project addressed the gender and race bias present in many collections by including female, Aboriginal and non-European convicts. Unlike the *Ironclad Sisterhood*, this project used bespoke commercial-grade image software to ensure the avatars matched the convict's description, including tattoos and scars. However, these images were not acknowledged as AI-generated, as the project team felt it was obvious they were AI-generated in the exhibition context.

### Findings and Insights

The project shifted public focus from secondary penal settlements, like Port Arthur, to other convict sites and convict lives outside these settlements, changing the historical understanding of convict lives. This shift introduced new ways to explore and interact with historical data, deepening the knowledge of convict networks within colonial society. For descendants of convicts and family historians, the project provided a more balanced perspective of the assignment system and the roles of convict and colonial networks.

Additionally, it shed light on convicts' roles in the violent dispossession of First Nations people from their firestick-farmed commons in the Tasmanian Midlands. Through AI-powered data analysis, the project enabled a thorough examination of the intricate connections between convicts and the colonial community, including their mobility, involvement in the development of labour markets, experiences within the law enforcement and criminal justice systems, and establishing the colonial class structure.

### **Analysis of all Projects**

The analysis of the three projects—*Ironclad Sisterhood*, *Convict Landscapes*, and *Making Crime Pay*—highlights the transformative potential of AI in historical research. The findings highlight the need for ethical guidelines in AI-driven historical research, reflecting broader discussions by Crymble, Milligan and Sherratt on ethical considerations and bias in digitisation.<sup>353 354 355</sup> They call for a balanced approach to AI's potential, ensuring innovation doesn't compromise historical authenticity.

#### Data Bias

Data bias in AI systems was a significant concern in the projects. In the *Ironclad Sisterhood*, AI-generated images of convict women were criticised for appearing too glamorous, likely due to biased training data or misuse prevention measures. *Making Crime Pay* addressed gender and racial biases by integrating diverse sources, aiming for more balanced representations. These projects emphasised the need to monitor AI algorithms to avoid misinterpreting historical data and mitigate unintended biases.

#### Ethical Use of AI

Ethical considerations were crucial, particularly in representing historical data with AI. The *Ironclad Sisterhood* addressed authenticity concerns by being transparent about AI use. *Landscape of Production and Punishment* highlighted the importance of monitoring AI to ensure unbiased historical and archaeological data interpretations. *Making Crime Pay* focused on creating sustainable digital assets and training AI on diverse sources for balanced representation, underscoring the role of ethics in AI deployment for historical research.

### Innovative Presentation of Historical Information

The projects used AI and digital technologies innovatively. The *Ironclad Sisterhood* created biographies and podcasts to explore convict women's lives. *Landscape of Production and Punishment* used AI-integrated LiDAR and GIS technologies for detailed historical maps. *Making Crime Pay* introduced the *Unshackled* exhibition with AI-generated convict avatars and digital storytelling, enhancing accessibility and engagement in historical research, aligning with Godfrey and Gibbs's recommendations.<sup>356 357</sup>

### **Conclusion**

Each project's unique AI application reinterprets Australia's convict heritage, demonstrating digital technologies' potential to offer new insights and challenge historical narratives. Similar projects worldwide leverage AI for new historical insights.<sup>358 359</sup> As AI becomes more prevalent, balancing its capabilities with guarding against bias and ensuring ethical stewardship of historical narratives is crucial. These projects contribute to vital debates in digital history, highlighting AI's transformative role in interpreting historical data for genealogical research.

#### 4.4 Future Prospects and Practical Implications of AI Integration in Genealogy:

As demonstrated in the previous sections, AI has significantly influenced genealogists' work practices and will continue to do so. At the end of data collection, a debriefing survey asked respondents to evaluate how AI and generative-AI enhanced their skills, confidence and ability to work more effectively, and whether they would use it again. The analysis revealed all participants were either currently using or planned to use AI in their future research, affirming that AI has effectively empowered them in their workflow. (Fig.57)

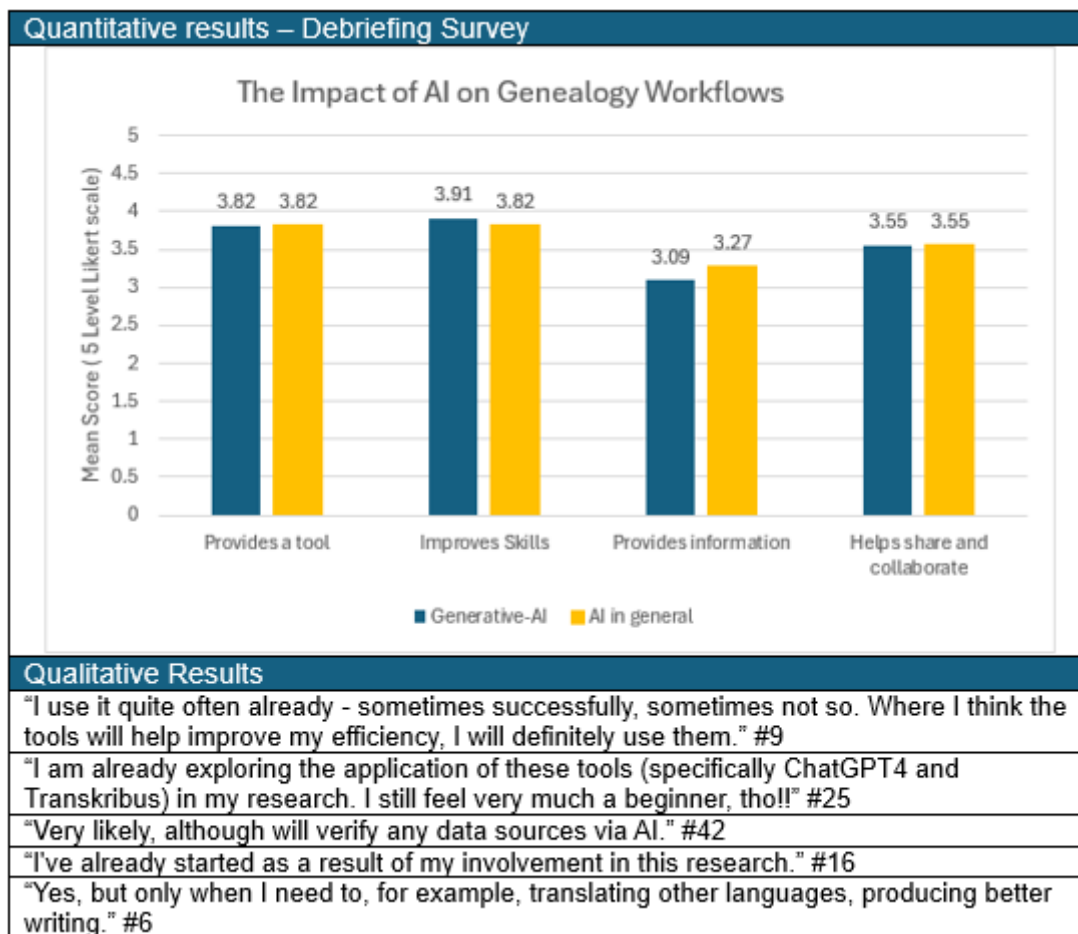


Figure 57 Evaluation of AI's ability to empower genealogists.

A closer look at the distribution of scores shows respondents perceive both types of AI positively as tools. (Fig.58) Generative-AI tends to be rated slightly higher than general AI across most fields, particularly in providing information and improving skills. The most significant variation in opinions is seen in sharing and collaboration, especially for generative-AI, where a notable outlier (rating of 1) exists. This suggests some disagreement among respondents or a lower awareness of generative AI's capabilities in this area. The baseline survey didn't capture feedback for this question, which could help identify reasons for this variation, an oversight that should be corrected in future investigations.

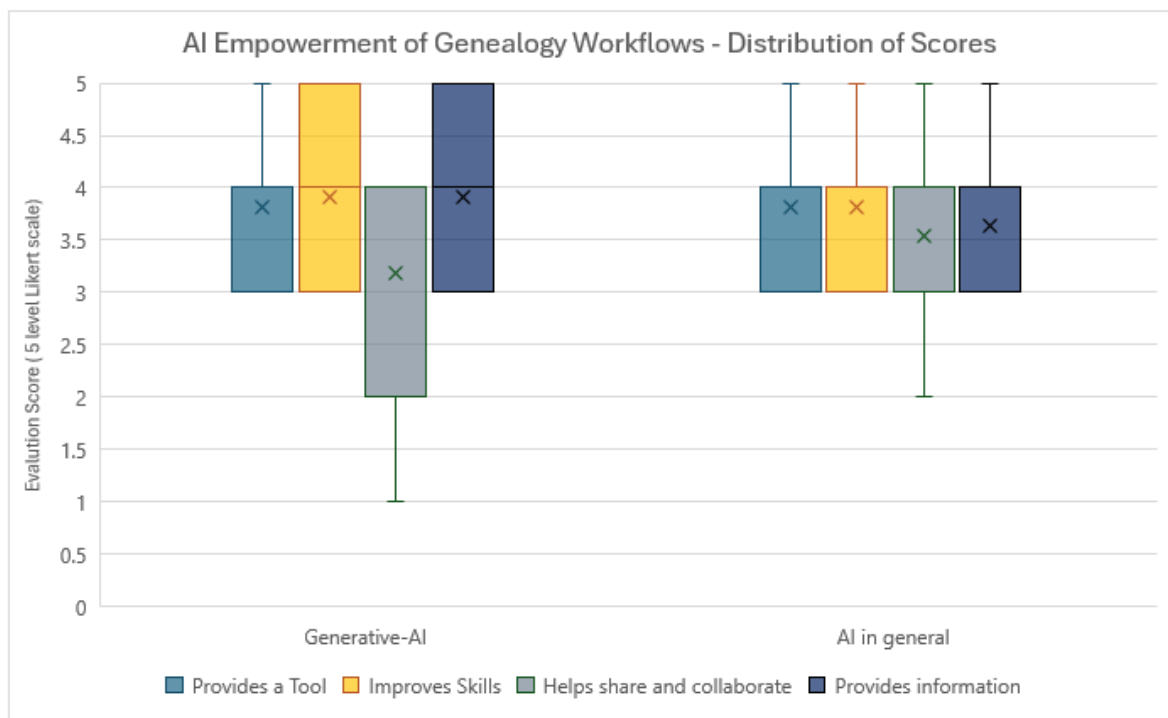


Figure 58 Distribution of scores for AI empowerment of genealogy workflows.

## Toward 2030

Respondents envision AI as an integral, well-accepted tool in genealogy, extensively used to interpret documents, compare data, and summarise DNA findings. (Table.17) .

Qualitative Results from Debriefing Survey
"I expect to be able to use AI more successfully in graphics and pictures, particularly when combining pictures with text. While I don't expect it to replace human effort in conducting genealogy research, I expect to use it to extract and summarise information from my own research and assist me in creating accurate narrative texts. I also expect it to help me in accurately transcribing and translating complex documents, thus improving my efficiency in undertaking time-consuming tasks." #9
"It will be an accepted research tool. Capabilities will be extended and improved. Users will be well educated in how to use the tools and their strengths and weaknesses." #25
"? Will have absorbed the big genealogical databases and perhaps helped clean up all the dodgy user trees, reduced the search labour involved, and perhaps shifted the emphasis from quantity to quality, understanding the historical context of ancestors' lives and building narratives around them." #11
"By then researchers will have sorted out how useful, or not, is AI. The youngsters of today will learn how to use AI, and it will be just another tool for geologists, as

Table 17 Predictions for the use of AI in genealogy in 2030.

They expect AI to enhance visual material creation, improve narrative accuracy, and streamline transcription and translation, making genealogy more accessible. AI is anticipated to evolve into a sophisticated research tool with known strengths and limitations, becoming a fundamental tool like DNA analysis. Respondents believe future genealogists, particularly the younger generation, will consider AI essential, integrated into major genealogical databases, improving data accuracy, reducing labour-intensive search processes, and emphasising quality and historical context.

## 4.5 Conclusion

The data from our cohort highlights a future where AI is poised to become an essential part of genealogy workflows, augmenting research with its growing capabilities. The depth and richness of the data gathered from respondents, with their experience in adapting to earlier technologies, have significantly contributed to understanding how AI can empower genealogists despite the cohort being relatively small. The findings are also significant for their timing, offering some of the earliest empirical data on AI's application in genealogy and establishing benchmark data for its integration into the field. Chapter 5 will assess the implications of these findings against the research objectives and offer recommendations for the successful integration of AI into genealogy.

## Chapter 5 – Conclusion

This research explored how AI, including generative-AI, can empower genealogists' workflows. A theoretical framework based on the literature review systematically addressed the research questions. This chapter summarises the key findings, highlights AI's significant contributions to genealogy workflows and concludes with personal reflections on the dissertation journey.

### 5.1 Generative-AI's Capabilities in Genealogy.

***RQ.1 What are the benefits and limitations of generative AI for genealogy purposes?***

***RQ.2 How can generative-AI be used in genealogy, how effective, efficient and accurate is it for genealogical purposes, and to what extent are genealogists aware of these tools?***

The literature review establishes that generative-AI augments knowledge workers' performance by enhancing speed, quantity, and quality.<sup>360</sup> However, its non-determinism and potential to generate non-factual information challenge its use in genealogy, which demands accuracy and reliability.<sup>361</sup> Clear, concise prompts and ample context can mitigate these challenges, allowing genealogists to benefit from AI's potential.<sup>362 363</sup>

Empirical evidence validated these insights: 72% of respondents said integrating generative-AI could improve research accuracy and depth, and 81.8% acknowledged the importance of verifying AI-generated information to ensure its integrity. It has limitations, notably record bias and hallucinations; however, genealogists are accustomed to working with inaccurate and biased information, making validation a routine part of their research process.

Generative-AI integrates well into genealogists' workflows for tasks like data extraction, summarising, translating and transcribing, and narrative construction. However, it struggled with tasks with more ambiguous data, such as creating a family tree, aligning with literature review insights.<sup>364 365 366</sup> Overall, it performed effectively to extremely effectively across various genealogical tasks.

The growing trend of genealogists incorporating generative-AI into their research practices is evident in the increased number of shared examples of AI's capabilities on social media, presentations and workshops.<sup>367</sup> Mollick emphasises the best results are achieved when AI and humans work collaboratively, a finding supported by this study, which showed AI augments rather than replaces genealogists' skills.<sup>368</sup> This collaborative approach enables

genealogists to use AI's speed and efficiency while applying their expertise to ensure accuracy and reliability.

As generative-AI continues to develop, improvements such as enhanced accuracy, better contextual understanding and bias mitigation will make these tools even more helpful for genealogical research.<sup>369</sup> The constant improvement of AI technologies highlights the importance of ongoing education and training to ensure genealogists remain well-informed about generative-AI's potential and how to safely integrate these tools into their workflows.

## **5.2 The Ethical Considerations of AI in Genealogy.**

### ***RQ.3 What ethical considerations and safeguards should be implemented when using AI, and what are the risks associated with not implementing them?***

Since its release in December 2022, the rapid adoption of generative-AI caught governments and organisations by surprise, leaving them scrambling to develop ethical frameworks for its responsible use.<sup>370</sup> The literature highlighted global concerns around data accuracy, privacy, AI's potential for bias and the need for transparent sourcing of training data. The study's findings echoed these concerns, with participants rating their concern level at 4.4/5.

Addressing these challenges requires collaborative efforts to build comprehensive ethical frameworks. Although respondents believe existing genealogy standards could be adapted for AI integration, they needed refinement. They identified five guidelines for the ethical AI integration:

1. Verification and validation of outputs,
2. Human oversight and curation,
3. Adherence to ethical and legal standards, particularly regarding privacy and sensitive data,
4. Clear documentation and acknowledgment of AI use in workflows
5. Development of comprehensive AI policies.

There was also a strong emphasis on the importance of community discussion and support, along with a call for genealogical organisations to take on an advocacy role and ensure genealogists' voices are heard in the broader ethical debates. Genealogy societies also play an essential role in supporting genealogists as they adapt to new technologies. By fostering continuous learning and ethical practice, these organisations ensure genealogists are well-equipped to integrate AI tools into their research.



Failure to implement safeguards can not only undermine the integrity and accuracy of genealogical research but can also risk the trust and reliability placed in their work by clients, family members and other genealogists.

### **5.3 The Broad Impact of AI on Genealogy.**

***RQ.4. How is AI integrated into current genetic genealogy processes, and what are the general benefits and challenges observed?***

***RQ.5. How is AI changing the availability, access and use of historical documents?***

***RQ.6. How is AI being used to re-interpret historical data to provide new insights for genealogical research?***

These questions examined the impact of various AI technologies beyond generative-AI, including DNA data analysis, document accessibility and reinterpretation of historical narratives.

The literature review revealed AI is fully integrated into genetic genealogy, facilitating every step of the DNA process.<sup>371</sup> Respondents used genetic data to supplement their research, with 82% having undertaken DNA testing. The study showed while AI enables powerful analytical tools for genetic genealogy, their use varied widely among respondents, indicating a need for ongoing education on integrating AI tools into DNA analysis.

AI technologies have also quietly changed how genealogists locate, access, and use records. The insights from previous technological changes provided in the literature aligned with the empirical evidence showing genealogists are integrating AI into their workflows.<sup>372</sup><sup>373</sup> Advances in AI algorithms for indexing, scanning and handwriting text recognition are making documents more digitally accessible.<sup>374</sup> The beta release of AI-enabled Natural Language searching could potentially revolutionise information retrieval, though current implementations received mixed feedback.<sup>375</sup> Respondents called for improved search functionality and resource coverage, issues expected to be resolved as the technology evolves.

The study found the respondents strongly prefer online research and use advanced AI capabilities such as predictive analysis, link prediction, and advanced filtering, often unaware these are sophisticated AI technologies. Despite a reliance on these capabilities, many respondents were not fully aware of the limitations and biases inherent in digital research, as highlighted in the literature review.<sup>376</sup><sup>377</sup> There is a clear need for ongoing education to equip genealogists with the skills to use these technologies effectively.<sup>378</sup>

This methodological shift is part of a broader move towards digital humanities, integrating digital technologies to augment traditional research methods. This was evident in the three real-world projects investigated during this research, which used AI-enabled databases, spatial technologies and AI story-telling techniques to provide new insights into the lives of convicts in colonial Australia. These projects faced limitations and challenges, particularly in balancing technological capabilities with ethical considerations and historical accuracy, reflecting consistent themes across both empirical evidence and the literature.<sup>379 380</sup>

#### **5.4 Contribution to Knowledge.**

The study revealed that generative-AI and AI can be integrated into genealogy workflows, empowering genealogists by enhancing their skills and boosting their confidence. It provides a snapshot of how an emerging technology is integrated into traditional genealogy workflows. The findings support the idea that AI enhances genealogists' performance and research quality when used collaboratively. These insights are especially valuable as the findings reflect the early stages of technology adoption, providing a baseline for future studies as the technology becomes more accepted. By addressing the initial impact and ethical considerations of generative-AI, it fills a significant gap in the literature. Moreover, the study contributes to broader discussions on the ethical use of AI and the need for robust standards to safeguard against potential risks while maximising benefits.

The research recommends ongoing education to help genealogists understand the benefits and limitations of the technology. Respondents felt the structured approach used in the Capability Analyses could form the basis for future educational courses or workshops.

## **5.5 Self-reflection**

Researching the use of AI and generative-AI in genealogy has been challenging and exciting. This rapidly evolving field required a steep learning curve and constant adaptation to remain current. Participating in some of the earliest courses and conferences addressing the use of generative-AI for genealogists has been particularly rewarding.

Initially, scholarly resources on generative-AI were scarce, which posed a challenge in the early stages of my research. As more resources emerged, they enriched my understanding and provided valuable validation of the research. These studies and articles helped bridge the knowledge gap and allowed for a more comprehensive exploration of AI in genealogy.

However, there are areas where the research could have been improved. Modifying some of the Capability Analyses could have provided better insights into practical genealogical applications of generative-AI. These adjustments could have enhanced the accuracy and relevance of the findings, offering a deeper understanding of AI's role in genealogy.

It was fascinating to witness the transformation among the study's respondents, from initial hesitation about incorporating AI into their workflows to becoming confident collaborators who understand how AI can enhance their research. They have become an informed group of advocates for generative-AI, spreading knowledge and encouraging others to explore its potential and benefiting the local genealogy community.

This evolution mirrors my experience; I've progressed from using generative-AI for isolated tasks to integrating it as a collaborative tool, improving my capabilities as a scholar and researcher. I also have a greater appreciation of the sophisticated AI technologies embedded in many aspects of genealogical research and their potential and limitations.

## 5.6 Future Directions

The next step for my research is developing education programs focused on the ethical and practical use of generative-AI in genealogy. These programs would equip genealogists with the skills and knowledge to use generative-AI, emphasising ethical considerations and the validation of AI-generated data.

This study has provided a foundational understanding of the integration of AI into genealogy. This is a developing field, with scope for further research in many areas to expand our knowledge of the potential of AI, including:

- Longitudinal Studies: Investigate the long-term impacts of AI integration into genealogy, tracking changes in research methodologies and outcomes over time, providing a deeper understanding of AI's sustained effects on genealogy.
- Broader Demographics: The research could include a more comprehensive range of genealogists from different backgrounds and skill levels. This could help determine how different groups use AI for genealogy research.
- More Diverse Range of Tasks: Examine the application of generative-AI across a broader range of genealogical tasks such as verifying source authenticity, translating/transcribing medieval documents, building a custom chatbot or integrating multi-media records like oral histories. This would provide a more holistic understanding of generative-AI's capabilities and limitations in genealogy.
- Ethical Framework Development: Develop and test ethical frameworks tailored explicitly to the use of AI in genealogy. This could involve multidisciplinary collaborations to balance technological advances with historical integrity and privacy concerns.

## **5.7 Conclusion**

The integration of AI into genealogical research represents a significant advance in genealogical methodologies. While not all genealogists may adopt generative-AI, those who do will find it transforms their workflows. This study demonstrated that effective AI integration enhances genealogists' efficiency and capabilities, directly improving research quality and depth to improve their research.

Education on AI's benefits and limitations is crucial, empowering genealogists to adapt to and excel in their work. The collaborative approach between AI and human expertise ensures routine tasks are streamlined while maintaining the accuracy and reliability of findings through critical human analysis.

As AI technology advances, its applications in genealogy will expand, offering sophisticated tools for data analysis and historical interpretation. This study provides a foundational understanding of AI's role in genealogy, emphasising the importance of ongoing education and ethical considerations. By embracing these advancements, genealogists can significantly enhance their research, empowering them to work more confidently and effectively.



## End Notes

---

- <sup>1</sup> MyHeritage. *AI & Genealogy: Harnessing The Power Of Artificial Intelligence For Family History Research - MyHeritage Knowledge Base*. <https://education.myheritage.com/article/ai-genealogy-harnessing-the-power-of-artificial-intelligence-for-family-history-research/> : accessed 22 May 2024.
- <sup>2</sup> Mollick, Ethan. (2024a) *Co-Intelligence: Living and Working with AI. [Kindle version]*. New York: Portfolio. <http://www.amazon.ca> : accessed 4 March 2024.
- <sup>3</sup> Walsh, Toby. (2023) *Faking It: Artificial Intelligence in a Human World*. Melbourne, Victoria: La Trobe University Press.
- <sup>4</sup> Marr, Bernard. *A Short History Of ChatGPT: How We Got To Where We Are Today*. <https://www.forbes.com/sites/bernardmarr/2023/05/19/a-short-history-of-chatgpt-how-we-got-to-where-we-are-today/?sh=4c7e7cfb674f> : accessed 21 May 2024.
- <sup>5</sup> Hu, Krystal. *ChatGPT sets record for fastest-growing user base - analyst note*. <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/> : accessed 21 May 2024.
- <sup>6</sup> Mollick (2024a) op. cit.
- <sup>7</sup> Open AI. *ChatGPT — Release Notes*. <https://help.openai.com/en/articles/6825453-chatgpt-release-notes> : accessed 11 June 2024.
- <sup>8</sup> Mollick (2024a) op. cit.
- <sup>9</sup> Dell'Acqua, Fabrizio, McFowland, Edward, Mollick, Ethan, et al. (2023) Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality. *Harvard Business School Technology & Operations Mgt. Unit Working Paper*. No. 24-013 pp. 1–58. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4573321](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4573321) : accessed 7 January 2024.
- <sup>10</sup> Hörnemalm, Adam. (2023) *ChatGPT as a Software Development Tool : The Future of Development*. Unpublished Master's Thesis, Umeå University. <https://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-209909> : accessed 22 October 2023.
- <sup>11</sup> Cooke, Lisa. (2023) Using ChatGPT or Bard for Genealogy. *Genealogy Gems*. [blog] 22 June. <https://lialouisecooke.com/2023/06/22/ai-chatbots-genealogy/> : accessed 21 May 2024.
- <sup>12</sup> Gabriel, John. (2023) *Has anyone tried ChatGPT for genealogy yet? - WikiTree G2G*. WikiTree G2G, <https://www.wikitree.com/g2g/1550687/has-anyone-tried-chatgpt-for-genealogy-yet> : accessed 22 May 2024.
- <sup>13</sup> Lee, Devon Noel. (2023) ChatGPT Failed? AI Struggles to Write Family History. *Family History Fanatics*. [blog] 12 June. <https://www.familyhistoryfanatics.com/chatgpt-writes-family-history> : accessed 22 May 2024.
- <sup>14</sup> Leeds, Dana. (2023) Using Artificial Intelligence in Genealogy. *Dana Leeds*. [blog] 24 October. <https://www.danaleeds.com/aigenealogy/> : accessed 21 May 2024.
- <sup>15</sup> Hoitink, Yvette. (2023) Five good and bad ways to use ChatGPT for Genealogy. *Dutch Genealogy*. [blog] 16 April. <https://www.dutchgenealogy.nl/five-good-and-bad-ways-to-use-chatgpt-for-genealogy/> : accessed 16 May 2024.
- <sup>16</sup> Whitford, Sara. (2023) WARNING! Be cautious using ChatGPT for genealogy research. *East Carolina Roots*. [blog] 27 March. <https://www.eastcarolinaroots.com/warning-be-cautious-using-chatgpt-for-genealogy-research/> : accessed 22 May 2024.
- <sup>17</sup> Little, Steve. (2024a) The Next Frontier: AI and Large Language Models Transforming Genealogy. *NGS Magazine*. 50 (1). pp. 20–24.
- <sup>18</sup> Ibid.
- <sup>19</sup> Mollick, Ethan. (2024b) Innovation through prompting. *One Useful Thing*. [blog] 22 April. <https://www.oneusefulthing.org/p/innovation-through-prompting> : accessed 23 April 2024.
- <sup>20</sup> Bettinger, Blaine T. (2023a) Leveraging Artificial Intelligence Tools for Genealogical Research. *Association of Professional Genealogists Quarterly*. December 2023 pp. 8–13.
- <sup>21</sup> Little (2024a) op. cit.
- <sup>22</sup> Little, Steve. *AI Genealogy Insights*. <https://aigenealogyinsights.com/author/digitalarchivist/> : accessed 4 February 2024.
- <sup>23</sup> Abdelhady, Ahmad M. & Davis, Christopher R. (2023) Plastic Surgery and Artificial Intelligence: How ChatGPT Improved Operation Note Accuracy, Time, and Education. *Mayo Clinic Proceedings: Digital Health*. 1 (3). pp. 299–308.

- 
- <https://www.sciencedirect.com/science/article/pii/S2949761223000470> : accessed 6 January 2024.
- <sup>24</sup> Lucy, Catherine. (2016) Research Trends & Emerging Technologies for Genealogists. *School of Information Student Research Journal*. 5 (2). pp. 1–23.  
<https://scholarworks.sjsu.edu/ischoolsrj/vol5/iss2/4> : accessed 9 December 2023.
- <sup>25</sup> Stallard, Matthew & de Groot, Jerome. (2020) “Things Are Coming Out That Are Questionable, We Never Knew About”: DNA and the New Family History. *Journal of Family History*. 45 (3). pp. 274–294. <https://doi.org/10.1177/0363199020906853> : accessed 10 January 2024.
- <sup>26</sup> Kellogg, Katherine, Lifshitz-Assaf, Hila, Randazzo, Steven, et al. (2024) Don't Expect Juniors to Teach Senior Professionals to Use Generative-AI: Emerging Technology Risks and Novice AI Risk Mitigation Tactics. *Harvard Business School Technology & Operations Mgt. Unit Working Paper*. 24–074 [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4857373](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4857373) : accessed 6 December 2024.
- <sup>27</sup> Creswell, John W. & Creswell, J. David. (2022) *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. [Kindle Version]. 6th ed. edition. Los Angeles: SAGE Publications. <https://www.amazon.ca> : accessed 15 January 2024.
- <sup>28</sup> Dell'Acqua, McFowland, Mollick, et al. (2023) op. cit.
- <sup>29</sup> Hörnemalm (2023) op. cit.
- <sup>30</sup> Denscombe, Martin. (2021) *The Good Research Guide: Research methods for small-scale social research projects*. 7th ed. New York: Open University Press.
- <sup>31</sup> Yin, Robert K. (2003) *Case Study Research: Design and Methods*. 3rd ed. Cambridge, USA: SAGE Publications, Inc. [https://iwansuharyanto.files.wordpress.com/2013/04/robert\\_k-\\_yin\\_case\\_study\\_research\\_design\\_and\\_mebookfi-org.pdf](https://iwansuharyanto.files.wordpress.com/2013/04/robert_k-_yin_case_study_research_design_and_mebookfi-org.pdf).
- <sup>32</sup> Mollick, Ethan. (2024c) What Apple's AI Tells Us: Experimental Models. *One Useful Thing*. [blog] 11 June. <https://www.oneusefulthing.org/p/what-apples-ai-tells-us-experimental> : accessed 11 June 2024.
- <sup>33</sup> Open AI. (2023b) *GPT-4 Technical Report*. Open AI. <https://cdn.openai.com/papers/gpt-4.pdf> : accessed 3 February 2024.
- <sup>34</sup> Mollick (2024c) op. cit.
- <sup>35</sup> Mollick, Ethan. (2024d) Doing Stuff with AI: Opinionated Mid-year Edition. *One Useful Thing*. [blog] 6 June. <https://www.oneusefulthing.org/p/doing-stuff-with-ai-opinionated-midyear> : accessed 6 June 2024.
- <sup>36</sup> Goodfellow, Ian, Bengio, Yoshua, & Courville, Aaron. (2016) *Deep Learning [pdf version]*. Cambridge, USA: The MIT Press. p.1 <https://www.deeplearningbook.org/> : accessed 23 January 2024.
- <sup>37</sup> Ibid., pp.3-4
- <sup>38</sup> Ibid., pp.5-8
- <sup>39</sup> Goodfellow, Bengio, & Courville (2016) op. cit.
- <sup>40</sup> Images: Chart. Venn diagram showing the connection between Artificial Intelligence, Machine Learning and Deep Learning. Bollard, Cheryl. December 2023.
- <sup>41</sup> Pillai, Abhilash. (2023) The Intersection of AI, Graph Neural Networks, and Genealogy: A New Era of Discovery. *Linked In*. <https://www.linkedin.com/pulse/intersection-ai-graph-neural-networks-genealogy-new-era-pillai> : accessed 8 October 2023.
- <sup>42</sup> Hörnemalm (2023) op. cit.
- <sup>43</sup> Ibid., pp.9-14
- <sup>44</sup> Mollick (2024c) op. cit.
- <sup>45</sup> Open AI. (2023a) *GPT-4 System Card*. Open AI. <https://cdn.openai.com/papers/gpt-4-system-card.pdf> : accessed 3 February 2024.
- <sup>46</sup> Ibid.
- <sup>47</sup> Ibid.
- <sup>48</sup> Ibid.
- <sup>49</sup> Hörnemalm (2023) op. cit. pp.9-20.
- <sup>50</sup> Manikandan, Bala. *Demystifying ChatGPT: A Deep Dive into Reinforcement Learning with Human Feedback*. <https://bmanikan.medium.com/demystifying-chatgpt-a-deep-dive-into-reinforcement-learning-with-human-feedback-1b695a770014> : accessed 3 February 2024.
- <sup>51</sup> Mollick (2024a) op. cit.
- <sup>52</sup> Kellogg, Lifshitz-Assaf, Randazzo, et al. (2024) op. cit.
- <sup>53</sup> Mollick (2024a) op. cit.



- 
- 54 Little (2024a) op. cit.
- 55 Kellogg, Lifshitz-Assaf, Randazzo, et al. (2024) op. cit.
- 56 Little (2024a) op. cit.
- 57 Mollick (2024b) op. cit.
- 58 Dell'Acqua, McFowland, Mollick, et al. (2023) op. cit.
- 59 Ibid., pp.17-19
- 60 Dell'Acqua, McFowland, Mollick, et al. (2023) op. cit.
- 61 Dell'Acqua, Fabrizio. (2022) *Falling Asleep at the Wheel: Human/AI Collaboration in a Field Experiment on HR Recruiters: Working Paper*. USA: Laboratory For Innovation Science, Harvard Business School.
- 62 Dell'Acqua, McFowland, Mollick, et al. (2023) op. cit. pp.20-25.
- 63 Ibid.
- 64 Dell'Acqua, McFowland, Mollick, et al. (2023) op. cit.
- 65 Ibid., p.23
- 66 Pillai (2023) op. cit.
- 67 Ancestry Team. (2019) How Ancestry® Is Preserving History One Record at a Time. *Ancestry Blog*. [blog] 26 August. <https://blogs.ancestry.com/cm/ancestry-imaging-center-what-does-it-take-to-create-an-image-on-ancestry/> : accessed 10 January 2024.
- 68 *Ancestry Hints*®. [https://support.ancestry.com/s/article/Ancestry-Hints?language=en\\_US](https://support.ancestry.com/s/article/Ancestry-Hints?language=en_US) : accessed 4 February 2024.
- 69 Bettinger (2023a) op. cit.
- 70 Ancestry Team. *Ancestry® to Apply Handwriting Recognition Artificial Intelligence to Create a Searchable Index of the 1950 U.S. Census*. <https://ancestry-prod2.adobecqms.net/content/ancestry-corp/en-us/blog/ancestry-apply-handwriting-recognition-artificial-intelligence-create-searchable-index-1950-us> : accessed 31 December 2023.
- 71 University of Exeter. *The Material Culture of Wills, England 1540-1790*. <https://sites.exeter.ac.uk/materialcultureofwills/about/> : accessed 9 April 2024.
- 72 Ibid.
- 73 Tanner, James L. (2024) 'Using Artificial Intelligence Tools to Expand your Genealogical Research Universe.' RootsTech 2024. Salt Lake City, Utah, USA <https://www.familysearch.org/en/rootstech/session/using-artificial-intelligence-tools-to-expand-your-genealogical-research-universe-part-1> : accessed 9 March 2024.
- 74 Gibbs, Frederick. (2016) New Forms of History: Critiquing Data and Its Representations. *The American Historian*. 7 pp. 31–36. <https://www.oah.org/tah/february-2/new-forms-of-history-critiquing-data-and-its-representations/> : accessed 7 March 2024.
- 75 Crymble, Adam. (2021) *Technology and the Historian: Transformations in the Digital Age*. [Kindle version]. Illinois: University of Illinois Press. [www.amazon.com.au](http://www.amazon.com.au) : accessed 17 December 2023.
- 76 Ibid.
- 77 Unknown photographer. Images: Photograph. Image of the card catalogue at the National Library of Australia, showing the transition from the card catalogue to computer/microfiche. August 1980. Canberra. Copy of photograph provided to author by the National Library of Australia April 2024.
- 78 Johnson, Elizabeth-Anne. (2017) *Toward a collaborative online framework for archival representation*. Unpublished Master's Thesis, University of Manitoba. [https://www.academia.edu/34550820/Toward\\_a\\_collaborative\\_online\\_framework\\_for\\_archival\\_representation](https://www.academia.edu/34550820/Toward_a_collaborative_online_framework_for_archival_representation) : accessed 28 December 2023.
- 79 Bell, Genevieve, Zafiroglu, Alex, Assaad, Zena, et al. (2021) *Custodians & Midwives: The Library of the Future*. Canberra: Australian National University. [https://cybernetics.anu.edu.au/assets/Custodians\\_and\\_Midwives\\_FINAL\\_web.pdf](https://cybernetics.anu.edu.au/assets/Custodians_and_Midwives_FINAL_web.pdf) : accessed 3 October 2023.
- 80 Jiang, Shuo, Luo, Jianxi, Hu, Jie, et al. (2022) Deep Learning for Technical Document Classification. *arXiv*. <https://arxiv.org/ftp/arxiv/papers/2106/2106.14269.pdf>.
- 81 Sherratt, Tim. (2019) 'Hacking Heritage: Understanding the limits of online access. In: Lewi, Hannah and Smith, Wally. eds.' *The Routledge International Handbook of New Digital Practices in Galleries, Libraries, Archives, Museums and Heritage Sites*. Routledge. <https://doi.org/10.4324/9780429506765> : accessed 8 March 2024.
- 82 Crymble (2021) op. cit.
- 83 Ibid.
- 84 Ibid.

- 
- <sup>85</sup> Sherratt (2019) op. cit.
- <sup>86</sup> Milligan, Ian. (2022) 'The Transformation of Historical Research in the Digital Age. In: Woolfe, Daniel. ed.' *Elements in Historical Theory and Practice*. Cambridge, UK: Cambridge University Press. pp. 1–38.
- <sup>87</sup> Ibid.
- <sup>88</sup> Ibid.
- <sup>89</sup> Godfrey, Barry, Homer, Caroline, Inwood, Kris, et al. (2021) Crime, Penal Transportation, and Digital Methodologies. *Journal of World History*. 32 (2). University of Hawai'i Press. pp. 241–260. <https://muse.jhu.edu/pub/5/article/794329> : accessed 27 February 2024.
- <sup>90</sup> Cowley, Trudy, Frost, Lucy, Inwood, Kris, et al. (2021) Reconstructing a Longitudinal Dataset for Tasmania. *Historical Life Course Studies*. 11 pp. 20–47. <https://hlcs.nl/article/view/10912> : accessed 2 March 2024.
- <sup>91</sup> Godfrey, Homer, Inwood, et al. (2021) op. cit.
- <sup>92</sup> Gibbs, Martin D., Tuffin, Richard L., Maxwell-Stewart, Hamish, et al. (2018) Landscapes of production and punishment: convict labour management on the Tasman Peninsula 1830–1877. *Antiquity*. 92 (362). <https://www.cambridge.org/core/journals/antiquity/article/landscapes-of-production-and-punishment-convict-labour-management-on-the-tasman-peninsula-18301877/A33CBE5287D38F307BFC526871276BB3> : accessed 1 March 2024.
- <sup>93</sup> Maxwell-Stewart, Hamish. (2022) *Making Crime Pay*. Australian Research Council Application. Project ID: LP220100118. Copy provided by author to Cheryl Bollard. October 2023.
- <sup>94</sup> Godfrey, Homer, Inwood, et al. (2021) op. cit.
- <sup>95</sup> Heraldry and Genealogy Society of Canberra. (1988) *Family History Beginners and Beyond*. 4th ed. Canberra ACT: Elect Printing.
- <sup>96</sup> Heraldry and Genealogy Society of Canberra. (2002) *Family History Beginners and Beyond*. 12th ed. Canberra ACT: Australian National University Printing Services.
- <sup>97</sup> Heraldry and Genealogy Society of Canberra. (2020) *Family History Beginners and Beyond*. 16th ed. Canberra ACT: Australian National University Printing Services.
- <sup>98</sup> Rogers, Everett M. (2003) *Diffusion of Innovations, 5th Edition*. 5th edition. New York: Free Press.
- <sup>99</sup> Ibid.
- <sup>100</sup> Ibid.
- <sup>101</sup> Kaminski, June. (2011) Diffusion of Innovation Theory. *Canadian Journal of Nursing Informatics*. 6 (2). <https://cjni.net/journal/?p=1444> : accessed 10 January 2024.
- <sup>102</sup> Duff, Wendy M. & Johnson, Catherine A. (2003) Where Is the List with All the Names? Information-Seeking Behavior of Genealogists. *The American Archivist*. 66 (1). pp. 79–95. <https://www.jstor.org/stable/40294218> : accessed 7 January 2024.
- <sup>103</sup> Duff, Wendy M. & Johnson, Catherine A. (2002) Accidentally Found on Purpose: Information-Seeking Behavior of Historians in Archives. *The Library Quarterly: Information, Community, Policy*. 72 (4). pp. 472–496. <https://www.jstor.org/stable/40039793> : accessed 7 January 2024.
- <sup>104</sup> Duff & Johnson (2003) op. cit.
- <sup>105</sup> Lucy (2016) op. cit.
- <sup>106</sup> Duff & Johnson (2002) op. cit.
- <sup>107</sup> Tibbo, Helen. (2003) Primarily History in America: How U.S. Historians Search for Primary Materials at the Dawn of the Digital Age. *The American Archivist*. 66 (1). pp. 9–50. <https://doi.org/10.17723/aarc.66.1.b1203701g718n74> : accessed 27 December 2023.
- <sup>108</sup> Yakel, Elizabeth. (2004) Seeking Information, Seeking Connections, Seeking Meaning: Genealogists and Family Historians. *Information Research*. 10 (1). <https://informationr.net/ir/10-1/paper205.html> : accessed 4 July 2024.
- <sup>109</sup> Lucy (2016) op. cit.
- <sup>110</sup> Ibid.
- <sup>111</sup> Bettinger (2023a) op. cit.
- <sup>112</sup> Ibid.
- <sup>113</sup> Lucy (2016) op. cit.
- <sup>114</sup> *Genealogy and Artificial Intelligence Facebook Group*. <https://www.facebook.com/groups/1255245945084761> : accessed 20 October 2023.
- <sup>115</sup> Little (2023a) op. cit.
- <sup>116</sup> Hoytink (2023) op. cit.
- <sup>117</sup> Leeds (2023) op. cit.

- 
- 118 Cain, David. (2023) Decoding the Human Blueprint: AI Revolutionizes Genetic Research. *Linked In*. <https://www.linkedin.com/pulse/genes-meet-machines-love-story-between-ai-genetic-research-david-cain/> : accessed 29 January 2024.
- 119 Ibid.
- 120 Images: Diagram. Chart showing the use of AI systems in genetic genealogy. Cheryl Bollard, author. January 2024.
- 121 Glynn, Claire L. (2022) Bridging Disciplines to Form a New One: The Emergence of Forensic Genetic Genealogy. *Genes*. 13 (8). pp. 1–16. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9407302/> : accessed 27 January 2024.
- 122 Ibid.
- 123 Dias, Raquel & Torkamani, Ali. (2019) Artificial intelligence in clinical and genomic diagnostics. *Genome Medicine*. 11 (1). p. 70. <https://doi.org/10.1186/s13073-019-0689-8> : accessed 8 December 2023.
- 124 Glynn (2022) op. cit.
- 125 Hanacek, N. *How Do Companies Measure DNA to Discover Your Ancestry?*. <https://www.nist.gov/how-do-you-measure-it/how-do-companies-measure-dna-discover-your-ancestry> : accessed 31 December 2023.
- 126 Holton, Graham S. (2019) *Tracing Your Ancestors Using DNA: A Guide for Family Historians*. [Kindle Version]. Pen and Sword. Chapter 1 <https://www.amazon.ca/> : accessed 3 February 2024.
- 127 Alharbi, Wardah S. & Rashid, Mamoon. (2022) A review of deep learning applications in human genomics using next-generation sequencing data. *Human Genomics*. 16 (1). pp. 5–20. <https://doi.org/10.1186/s40246-022-00396-x> : accessed 30 December 2023.
- 128 AncestryDNA®. *Genotype*. <https://www.ancestry.com/c/dna-learning-hub/genotype> : accessed 28 January 2024.
- 129 Holton (2019) op. cit. chap. Chapter.1.
- 130 Ibid., Chapter 1
- 131 Stankus, Tony. (2020) Reviews of Science for Science Librarians: Direct-to-Consumer DNA Testing for Ancestry as a Complement to Traditional Genealogical Methods. *Science & Technology Libraries*. 39 (3). pp. 227–243. <https://doi.org/10.1080/0194262X.2020.1758285> : accessed 27 January 2024.
- 132 Adrion, Jeffrey, Lang, Jenna, Noto, Keith, et al. (2023) *Ethnicity Estimate 2023 White Paper*. Ancestry.com. <https://www.ancestrycdn.com/support/us/2023/09/Ethnicity2023whitepaper.pdf> : accessed 5 February 2024.
- 133 Hu, Rui, Maier, Paul, Runfeldt, Goran, et al. (2021) *Family Finder Matching 5.0: Matching Algorithm and Relationship Estimation White Paper*. FamilyTreeDNA. [https://blog.familytreedna.com/wp-content/uploads/2021/08/Family\\_Finder\\_Matching\\_WhitePaper.pdf](https://blog.familytreedna.com/wp-content/uploads/2021/08/Family_Finder_Matching_WhitePaper.pdf) : accessed 4 February 2024.
- 134 Bettinger, Blaine T. (2016) *The Family Tree Guide to DNA Testing and Genetic Genealogy: How to Harness the Power of DNA to Advance Your Family Tree Research*. London: Penguin.
- 135 Hu, Maier, Runfeldt, et al. (2021) op. cit.
- 136 Ball, Catherine A, Barber, Mathew J, Byrnes, Jake, et al. (2016) *AncestryDNA Matching White Paper: Discovering genetic matches across a massive, expanding genetic database*. Ancestry.com. <https://www.ancestry.com/cs/dna-help/matches/whitepaper> : accessed 4 February 2024.
- 137 Ancestry Team. *How SideView™ Technology Splits Your DNA Results by Parent*. <https://support.ancestry.com.au/s/article/SideView-Technology> : accessed 28 January 2024.
- 138 Bettinger (2016) op. cit.
- 139 Stallard & de Groot (2020) op. cit.
- 140 Holton, Graham S. (2020) Discovering unknown medieval descents : a genetic approach – medieval genealogy for the masses. *Foundation*. 12 pp. 2–15. [https://strathprints.strath.ac.uk/73099/1/Holton\\_Foundations\\_2020\\_Discovering\\_unknown\\_medieval\\_descents\\_a\\_genetic\\_approach\\_medieval\\_genealogy\\_for\\_the\\_masses.pdf](https://strathprints.strath.ac.uk/73099/1/Holton_Foundations_2020_Discovering_unknown_medieval_descents_a_genetic_approach_medieval_genealogy_for_the_masses.pdf) : accessed 4 February 2024.
- 141 Christofides, Emily & O'Doherty, Kieran. (2016) Company disclosure and consumer perceptions of the privacy implications of direct-to-consumer genetic testing. *New Genetics and Society*. 35 (2). pp. 101–123. <https://www.tandfonline.com/doi/full/10.1080/14636778.2016.1162092> : accessed 29 January 2024.

- 
- 142 Parikh, Ravi B., Teeple, Stephanie, & Navathe, Amol S. (2019) Addressing Bias in Artificial Intelligence in Health Care. *JAMA*. 322 (24). pp. 2377–2378. <https://doi.org/10.1001/jama.2019.18058> : accessed 29 January 2024.
- 143 Ménard, Timothé. (2022) Good quality practices for artificial intelligence in genetics. *European Journal of Human Genetics*. 30 (9). pp. 993–995. <https://www.nature.com/articles/s41431-022-01061-6> : accessed 8 December 2023.
- 144 Royal, Charmaine D., Novembre, John, Fullerton, Stephanie M., et al. (2010) Inferring Genetic Ancestry: Opportunities, Challenges, and Implications. *American Journal of Human Genetics*. 86 (5).: pp.661-668 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2869013/> : accessed 30 January 2024.
- 145 Genetic Genealogy Standards Committee. *Genetic Genealogy Standards*. <http://www.geneticgenealogystandards.com/> : accessed 5 April 2024.
- 146 ISOGG Wiki. *Ethics, Guidelines and Standards*. [https://isogg.org/wiki/Ethics,\\_guidelines\\_and\\_standards](https://isogg.org/wiki/Ethics,_guidelines_and_standards) : accessed 2 February 2024.
- 147 Kranzberg, Melvin. (1986) Technology and History: “Kranzberg’s Laws.” *Technology and Culture*. 27 (3). pp. 544–560. <https://www.jstor.org/stable/3105385> : accessed 2 February 2024.
- 148 Mollick, Ethan. (2024e) What Can be Done in 59 Seconds: An Opportunity (and a Crisis). *One Useful Thing*. [blog] 31 January. [https://www.oneusefulthing.org/p/what-can-be-done-in-59-seconds-an?utm\\_campaign=post&fbclid=IwAR1hZDJdaCDG0mH3FRmf8NqY4P\\_F4ijuCMc9mlaiAGdupSU70XRpyZuCJvc](https://www.oneusefulthing.org/p/what-can-be-done-in-59-seconds-an?utm_campaign=post&fbclid=IwAR1hZDJdaCDG0mH3FRmf8NqY4P_F4ijuCMc9mlaiAGdupSU70XRpyZuCJvc) : accessed 2 February 2024.
- 149 Ibid.
- 150 Fjeld, Jessica, Achten, Nele, Hilligoss, Hannah, et al. (2020) *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI*. Rochester, NY: The Berkman Klein Center for Internet & Society Research. <https://papers.ssrn.com/abstract=3518482> : accessed 31 January 2024.
- 151 Ibid.
- 152 Chatila, Raja, Firth-Butterfield, Kay, & Havens, John C. (2018) *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems Version 2*. Los Angeles: University of Southern California. <https://apps.dtic.mil/sti/citations/AD1170922> : accessed 1 February 2024.
- 153 Edwards, Lilian. (2022) *The EU AI Act proposal: a summary of its significance and scope*. London, UK: Ada Lovelace Institute. pp.15-18 <https://www.adalovelaceinstitute.org/resource/eu-ai-act-explainer/>.
- 154 Li, Fei-Fei. (2023) *The Worlds I See: Curiosity, Exploration, and Discovery at the Dawn of AI. [Kindle Version]*. New York: Flatiron Books. <http://amazon.com.au> : accessed 1 February 2024.
- 155 Mollick (2024a) op. cit.
- 156 Ibid.
- 157 MacEntee, Thomas. *A Genealogist’s Guide to Artificial Intelligence*. <https://familytreemagazine.com/resources/software/ai-and-genealogy/> : accessed 28 January 2024.
- 158 Cowgill, Bo, Dell’Acqua, Fabrizio, Deng, Samuel, et al. (2020) ‘Biased Programmers? Or Biased Data? A Field Experiment in Operationalizing AI Ethics.’ In: *Proceedings of the 21st ACM Conference on Economics and Computation*, 21st Conference on Economics and Computation. New York, NY, USA: Association for Computing Machinery. pp. 679–681. <https://dl.acm.org/doi/10.1145/3391403.3399545> : accessed 4 February 2024.
- 159 Bettinger (2023a) op. cit. p.10.
- 160 Ibid.
- 161 Ibid.
- 162 Kaebnick, Gregory E., Magnus, David Christopher, Kao, Audiey, et al. (2023) Editors’ Statement on the Responsible Use of Generative AI Technologies in Scholarly Journal Publishing. *Hastings Center Report*. 53 (5). pp. 3–6. <https://onlinelibrary.wiley.com/doi/abs/10.1002/hast.1507> : accessed 2 May 2024.
- 163 Hosseini, Mohammad, Holmes, Kristi, & Resnick, David. (2023) The ethics of disclosing the use of artificial intelligence tools in writing scholarly manuscripts - Mohammad Hosseini, David B Resnik, Kristi Holmes, 2023. *Research Ethics*. 19 (4). p. pp.449-465. <https://journals-sagepub-com.proxy.lib.strath.ac.uk/doi/10.1177/17470161231180449> : accessed 2 May 2024.

- 
- 164 Little, Steve. (2023b) 'Empowering Genealogists with Artificial Intelligence: Syllabus Week 2.' National Genealogical Society (US) course. USA.
- 165 Little, Steve. (2024b) Avoiding Extremism: The Use and Disclosure of AI in Genealogy. *AI Genealogy Insights*. [blog] 9 June. <https://aigenealogyinsights.com/2024/06/09/avoiding-extremism-the-use-and-disclosure-of-ai-in-genealogy/> : accessed 10 June 2024.
- 166 Mollick (2024a) op. cit.
- 167 MacEntee (2023) op. cit.
- 168 Bettinger (2023a) op. cit.
- 169 Little (2023b) op. cit.
- 170 Board for Certification of Genealogists. *Genealogist's Code of Ethics*. <https://bcgcertification.org/ethics-standards/code/> : accessed 2 February 2024.
- 171 Register of Qualified Genealogists. *Professional Code*. <https://www.qualifiedgenealogists.org/the-professional-code> : accessed 19 February 2024.
- 172 Australasian Association of Genealogists and Record Agents. *AAGRA Code of Ethics*. <https://www.aagra.asn.au/aagra-code-of-ethics/> : accessed 2 February 2024.
- 173 Chatila, Firth-Butterfield, & Havens (2018) op. cit.
- 174 Bettinger, Blaine T. (2023b) *This group is at the forefront of genealogy and AI*, Genealogy and Artificial Intelligence, [Facebook post]. 19 November. <https://www.facebook.com/groups/genealogyandai/permalink/1414732652469422> : accessed 26 January 2024.
- 175 Benn, Ashley. (2023) *I am one of the many who is conflicted on whether or not to use AI to illustrate our genealogy*, Genealogy and Artificial Intelligence, [Facebook post]. 19 November. <https://www.facebook.com/groups/genealogyandai/permalink/1414556692487018> : accessed 26 January 2024.
- 176 Creswell & Creswell (2022) op. cit.
- 177 Ibid.
- 178 Hörnemalm (2023) op. cit.
- 179 Dell'Acqua, McFowland, Mollick, et al. (2023) op. cit.
- 180 Creswell & Creswell (2022) op. cit.
- 181 Ibid.
- 182 Mollick (2024a) op. cit.
- 183 Rogers (2003) op. cit.
- 184 Little (2024a) op. cit.
- 185 Mollick (2024a) op. cit.
- 186 Dell'Acqua, McFowland, Mollick, et al. (2023) op. cit.
- 187 Rogers (2003) op. cit.
- 188 Dell'Acqua, McFowland, Mollick, et al. (2023) op. cit.
- 189 Rogers (2003) op. cit.
- 190 Creswell & Creswell (2022) op. cit. chap.10.
- 191 Creswell & Creswell (2022) op. cit.
- 192 Ibid., chap.10
- 193 Images: Diagram. Steps used for inductive coding of quantitative data. Cheryl Bollard, author. June 2024.
- 194 Lucy (2016) op. cit.
- 195 Yin (2003) op. cit. pp1-10.
- 196 Ibid., pp.1-10
- 197 Denscombe (2021) op. cit. pp.229-231.
- 198 Ibid., p.231
- 199 Rogers (2003) op. cit.
- 200 Creswell & Creswell (2022) op. cit.
- 201 Mollick (2024b) op. cit.
- 202 Mollick (2024d) op. cit.
- 203 Open AI (2023a) op. cit.
- 204 Open AI. *Introducing GPT-4o and more tools to ChatGPT free users*. <https://openai.com/index/gpt-4o-and-more-tools-to-chatgpt-free/> : accessed 27 May 2024.
- 205 Duff & Johnson (2003) op. cit.
- 206 Tanner (2024) op. cit.
- 207 Little (2024a) op. cit.

- 
- 208 Testamentary records. England. 18 March 1720. ARNOLD, Coe. Will. Norwich Consistory Court Probate Records: Will Registers. NCC will register Archer 1. Norfolk Record Office, Norfolk, England. <https://nrocatalogue.norfolk.gov.uk/arnold-coe-beer-brewer-of-lowestoft-suffolk> : accessed 29 January 2024.
- 209 Open AI. (2024c) *ChatGPT (GPT-4, custom GPT Genealogy Summariser May 2024 version)*. Response to chat with C. Bollard on the summary of Coe ARNOLD's will. 2 June 2024. See Appendix H for prompt used and output generated.
- 210 Testamentary records. England. 18 March 1720. ARNOLD, Coe. Will. Norwich Consistory Court Probate Records: Will Registers. NCC will register Archer 1. Norfolk Record Office, Norfolk, England. <https://nrocatalogue.norfolk.gov.uk/arnold-coe-beer-brewer-of-lowestoft-suffolk> : accessed 1 February 2024
- 211 Open AI. (2024d) *ChatGPT (GPT-4, May 13 Version)* Response to chat with C. Bollard on the extraction of data by person. 2-6 June 2024. See Appendix H for prompt and output generated.
- 212 Open AI. (2024e) *ChatGPT (GPT-4, May 13 Version)* Response to chat with C. Bollard on the extraction of data by property. 2-6 June 2024. See Appendix H for prompt and output generated.
- 213 Hoitink, Yvette. (2024) 'Handwritten Text Recognition: Harnessing the Power of Artificial Intelligence to Transcribe and Search.' In: *Session 203 Syllabus , Expanding Possibilities*. Virginia, USA: National Genealogical Society. <https://www.ngsgenealogy.org/> : accessed 18 May 2024.
- 214 Births index (CR) Netherlands. Noord-Holland. 22 March 1841. AMEN, Wilhelmus Johannes Franciscus Henricus. 6 January 2024. <https://www.familysearch.org/ark:/61903/1:1:Q2H6-PQFN> : accessed 12 January 2024.
- 215 FamilySearch. *Wiki: Netherlands Civil Registration Records*. [https://www.familysearch.org/en/wiki/Netherlands\\_Civil\\_Registration\\_Record](https://www.familysearch.org/en/wiki/Netherlands_Civil_Registration_Record) : accessed 27 May 2024.a.
- 216 Open AI. (2024f) *ChatGPT (GPT-4, May 13 Version)* Response to chat with C. Bollard on translation of Dutch birth record. 28 May 2024. See Appendix H for prompt and output generated.
- 217 Little (2024a) op. cit.
- 218 Mollick (2024a) op. cit.
- 219 Obituaries. (1938) *GoulburnValley Stock and Property Journal*. 13 July. Victoria WOODS, J.T. p. 8. <https://trove.nla.gov.au/newspaper/article/219771008> : accessed 7 January 2024.
- 220 Open AI. (2024g) *ChatGPT (GPT-4, May 13 Version)* . Response to chat with C. Bollard on a J.T. WOOD narrative. 29 May 2024. See Appendix H for prompt used and output generated.
- 221 Facebook. *Genealogy and Artificial Intelligence (AI) | Facebook Group*. <https://www.facebook.com/groups/genealogyandai> : accessed 23 April 2024.
- 222 FamilySearch. *RootsTech Videos 2023-2024*. <https://www.familysearch.org/en/rootstech/> : accessed 21 April 2024.
- 223 Little, Steve. (2024c) *Congrats to Dana Stewart Leeds and Thomas MacEntee for Top 10 finishes with AI topics for March 2024 Legacy Family Tree Webinars* Genealogy and Artificial Intelligence (AI), [Facebook post]. 2 April. <https://www.facebook.com/groups/genealogyandai/permalink/1489472768328743> : accessed 21 April 2024.
- 224 Rogers (2003) op. cit.
- 225 Dell'Acqua, McFowland, Mollick, et al. (2023) op. cit.
- 226 Hörnemalm (2023) op. cit.
- 227 Survey Monkey. *Pearson vs Spearman correlations: practical applications*. <https://www.surveymonkey.com/market-research/resources/pearson-correlation-vs-spearman-correlation/> : accessed 6 April 2024.
- 228 Little (2024a) op. cit.
- 229 Bettinger (2023a) op. cit.
- 230 Little (2024a) op. cit.
- 231 Bettinger (2023a) op. cit.
- 232 Kellogg, Lifshitz-Assaf, Randazzo, et al. (2024) op. cit.
- 233 Mollick (2024a) op. cit.
- 234 Dell'Acqua, McFowland, Mollick, et al. (2023) op. cit.
- 235 Mollick (2024a) op. cit.
- 236 Little (2023b) op. cit.

- 237 Little, Steve. (2023c) New Use Case: Handwriting Recognition to Structured Data. *AI Genealogy Insights*. [blog] 9 October. <https://aigenealogyinsights.com/2023/10/09/new-use-case-handwriting-recognition-to-structured-data/> : accessed 17 October 2023.
- 238 Kellogg, Lifshitz-Assaf, Randazzo, et al. (2024) op. cit.
- 239 Mollick (2024a) op. cit.
- 240 Kellogg, Lifshitz-Assaf, Randazzo, et al. (2024) op. cit.
- 241 Dell'Acqua, McFowland, Mollick, et al. (2023) op. cit.
- 242 Testamentary records. England. 18 March 1720. ARNOLD, Coe. Will. Norwich Consistory Court Probate Records: Will Registers. NCC will register Archer 1. Norfolk Record Office, Norfolk, England. <https://nrcatalogue.norfolk.gov.uk/arnold-coe-beer-brewer-of-lowestoft-suffolk> : accessed 29 January 2024.
- 243 Open AI (2024d) op. cit.
- 244 Open AI. (2024h) *ChatGPT (GPT-3.5, Feb 2024 Version) Response to chat with #9 on the creation of a family tree. February 2024*. See Appendix H for prompt used and output generated.
- 245 Open AI. (2024i) *ChatGPT (GPT-4, May 13 Version) Response to chat with C. Bollard on the creation of a family tree. 29 May 2024*. See Appendix H for prompt used and output generated.
- 246 Rogers (2003) op. cit.
- 247 Lucy (2016) op. cit.
- 248 Little (2024a) op. cit.
- 249 Bettinger (2023a) op. cit.
- 250 MacEntee (2023) op. cit.
- 251 Fjeld, Achten, Hilligoss, et al. (2020) op. cit.
- 252 Li (2023) op. cit.
- 253 Mollick (2024a) op. cit.
- 254 MacEntee (2023) op. cit.
- 255 Bettinger (2023a) op. cit.
- 256 Little (2024a) op. cit.
- 257 Royal, Novembre, Fullerton, et al. (2010) op. cit.
- 258 Rogers (2003) op. cit.
- 259 Australasian Association of Genealogists and Record Agents (2015) op. cit.
- 260 Board for Certification of Genealogists (2020) op. cit.
- 261 Register of Qualified Genealogists (n.d.) op. cit.
- 262 Bettinger (2023a) op. cit.
- 263 Bettinger (2023b) op. cit.
- 264 Open AI. (2024j) *ChatGPT (GPT-4, March Version) . Response to chat with C. Bollard on the creation of an ethical framework. 29 May 2024*. See Appendix H for prompt used and output generated.
- 265 Bettinger (2023a) op. cit.
- 266 Little (2024a) op. cit.
- 267 Mollick (2024a) op. cit.
- 268 Cowgill, Dell'Acqua, Deng, et al. (2020) op. cit.
- 269 Mollick (2024a) op. cit.
- 270 Ibid.
- 271 Dell'Acqua (2022) op. cit.
- 272 Dell'Acqua, McFowland, Mollick, et al. (2023) op. cit.
- 273 Mollick (2024a) op. cit. p.44.
- 274 Cain (2023) op. cit.
- 275 Christofides & O'Doherty (2016) op. cit.
- 276 Hu, Maier, Runfeldt, et al. (2021) op. cit.
- 277 Ball, Barber, Byrnes, et al. (2016) op. cit.
- 278 Survey Monkey (2024) op. cit.
- 279 MacEntee (2023) op. cit.
- 280 Southard, Diahhan. (2021) Understanding AncestryDNA Genetic Communities. *Your DNA Guide - Diahhan Southard*. [blog] 1 February. <https://www.yourdnaguide.com/ydgblog/ancestrydna-genetic-communities> : accessed 31 May 2024.
- 281 My Heritage. *Autoclusters - How to Use Autoclusters on MyHeritage*. <https://education.myheritage.com/article/autoclusters-for-dna-matches/> : accessed 31 May 2024.
- 282 MyHeritage (2024) op. cit.

- 
- 283 Stallard & de Groot (2020) op. cit.
- 284 Holton (2019) op. cit. chap.1.
- 285 Bettinger (2016) op. cit.
- 286 Hanacek (n.d.) op. cit.
- 287 Holton (2019) op. cit. chap.4.
- 288 Stankus (2020) op. cit.
- 289 Bettinger (2016) op. cit. pp.35-50.
- 290 Christofides & O'Doherty (2016) op. cit.
- 291 Holton (2019) op. cit. chap.2.
- 292 Genetic Genealogy Standards Committee (2015) op. cit.
- 293 ISOGG Wiki (n.d.) op. cit.
- 294 Pillai (2023) op. cit.
- 295 Bell, Zafiroglu, Assaad, et al. (2021) op. cit.
- 296 Tanner (2024) op. cit.
- 297 Bettinger (2023a) op. cit.
- 298 Little (2023c) op. cit.
- 299 Hoitink (2024) op. cit.
- 300 Bell, Zafiroglu, Assaad, et al. (2021) op. cit.
- 301 Johnson (2017) op. cit.
- 302 Erica. *Introducing AI Record Finder™, the World's First AI Chat-Based Search Engine for Historical Records*. <https://blog.myheritage.com/2023/12/introducing-ai-record-finder-the-worlds-first-ai-chat-based-search-engine-for-historical-records/> : accessed 31 December 2023.
- 303 FamilySearch. *Full-Text Search*. <https://www.familysearch.org/search/full-text> : accessed 10 April 2024.
- 304 Creswell & Creswell (2022) op. cit.
- 305 My Heritage. (2024) *AI Record Finder. ( May 2024 Version) Response to a search request by C. Bollard for information on John DRAKE. 16 May 2024*. See Appendix F for prompt used and output generated.
- 306 Creswell & Creswell (2022) op. cit.
- 307 Open AI (2023a) op. cit.
- 308 (N.d.) op. cit.
- 309 Lucy (2016) op. cit.
- 310 Pillai (2023) op. cit.
- 311 Ibid.
- 312 Duff & Johnson (2003) op. cit.
- 313 Duff & Johnson (2002) op. cit.
- 314 Open AI. (2024k) *ChatGPT (GPT-4, May 13 Version) Response to chat with C. Bollard on the creation of a research plan. 29 May 2024*. See Appendix H for prompt and output generated.
- 315 Lucy (2016) op. cit.
- 316 Ibid.
- 317 Little (2023a) op. cit.
- 318 Johnson (2017) op. cit.
- 319 Sherratt (2019) op. cit.
- 320 Milligan (2022) op. cit.
- 321 Gibbs, Tuffin, Maxwell-Stewart, et al. (2018) op. cit.
- 322 Maxwell-Stewart (2022) op. cit.
- 323 Sherratt (2019) op. cit.
- 324 Crymble (2021) op. cit.
- 325 Gibbs (2016) op. cit.
- 326 Milligan (2022) op. cit.
- 327 Open AI (2023b) op. cit.
- 328 Society of Australian Genealogists. *Ironclad Sisterhood: Convict Women in Australia*. <https://ironcladsisterhood.sag.org.au/> : accessed 9 March 2024.
- 329 Benn (2023) op. cit.
- 330 Bettinger (2023b) op. cit.
- 331 Crymble (2021) op. cit.
- 332 Godfrey, Homer, Inwood, et al. (2021) op. cit.
- 333 Gibbs (2016) op. cit.



- 
- 334 Milligan (2022) op. cit.
- 335 Cowley, Frost, Inwood, et al. (2021) op. cit.
- 336 Ibid.
- 337 Godfrey, Homer, Inwood, et al. (2021) op. cit.
- 338 Little (2024a) op. cit.
- 339 Maxwell-Stewart (2022) op. cit.
- 340 Cowley, Frost, Inwood, et al. (2021) op. cit.
- 341 Tuffin, Richard, Gibbs, Martin, Roberts, David, et al. (2018) Landscapes of Production and Punishment: Convict labour in the Australian context. *Journal of Social Archaeology*. 18 (1). pp. 50–76. <https://journals.sagepub.com/doi/epub/10.1177/1469605317748387> : accessed 21 October 2023.
- 342 Images: Photograph. Port Arthur Convict Quarries. 12 October 2023. Port Arthur, Tasmania. Jeffrey Bollard, photographer. 2023. a. Private collection of author, Canberra.
- 343 Cowley, Frost, Inwood, et al. (2021) op. cit.
- 344 Gibbs, Tuffin, Maxwell-Stewart, et al. (2018) op. cit.
- 345 Milligan (2022) op. cit.
- 346 Crymble (2021) op. cit.
- 347 Milligan (2022) op. cit.
- 348 Maxwell-Stewart (2022) op. cit.
- 349 University of New England. *World-first digital project gives unique insights into convict life*. <https://www.une.edu.au/connect/news/2023/11/world-first-digital-project-gives-unique-insights-into-convict-life> : accessed 27 February 2024.
- 350 Images: Photograph. Convict Memorial Tower. 6 October 2023. Hobart Penitentiary Chapel, Hobart, Tasmania. Jeffrey Bollard, photographer. 2023. b. Private collection of author, Canberra.
- 351 University of New England (2023) op. cit.
- 352 Maxwell-Stewart (2022) op. cit.
- 353 Crymble (2021) op. cit.
- 354 Milligan (2022) op. cit.
- 355 Sherratt (2019) op. cit.
- 356 Godfrey, Homer, Inwood, et al. (2021) op. cit.
- 357 Gibbs (2016) op. cit.
- 358 Ahnert, Ruth. (2023) *The Living with Machines Report [pdf version]*. Arts and Humanities Research Council. [https://livingwithmachines.ac.uk/wp-content/uploads/2023/11/ATI007%E2%80%9494Living-With-Machines\\_FINAL-Fixed.pdf](https://livingwithmachines.ac.uk/wp-content/uploads/2023/11/ATI007%E2%80%9494Living-With-Machines_FINAL-Fixed.pdf) : accessed 30 December 2023.
- 359 Open Culture. (n.d.) The Venice Time Machine: 1,000 Years of Venice’s History Gets Digitally Preserved with Artificial Intelligence and Big Data. *Open Culture*. [blog] <https://www.openculture.com/2019/03/the-venice-time-machine.html> : accessed 3 June 2024.
- 360 Dell’Acqua, McFowland, Mollick, et al. (2023) op. cit.
- 361 Bettinger (2023a) op. cit.
- 362 Little (2024a) op. cit.
- 363 Mollick (2024a) op. cit.
- 364 Dell’Acqua, McFowland, Mollick, et al. (2023) op. cit.
- 365 Kellogg, Lifshitz-Assaf, Randazzo, et al. (2024) op. cit.
- 366 Mollick (2024a) op. cit.
- 367 Facebook (2024) op. cit.
- 368 Mollick (2024a) op. cit.
- 369 Open AI. *Research*. <https://openai.com/research/> : accessed 19 May 2024.
- 370 Fjeld, Achten, Hilligoss, et al. (2020) op. cit.
- 371 Stallard & de Groot (2020) op. cit.
- 372 Duff & Johnson (2003) op. cit.
- 373 Lucy (2016) op. cit.
- 374 Hoytink (2024) op. cit.
- 375 Bell, Zafiroglu, Assaad, et al. (2021) op. cit.
- 376 Milligan (2022) op. cit.
- 377 Crymble (2021) op. cit.
- 378 Milligan (2022) op. cit.
- 379 Gibbs (2016) op. cit.
- 380 Godfrey, Homer, Inwood, et al. (2021) op. cit.



## Bibliography

### Primary Sources

Births index (CR) Netherlands. Noord-Holland. 22 March 1841. AMEN, Whilhelmus Johannes Franciscus Henricus. 6 January 2024. <https://www.familysearch.org/ark:/61903/1:1:Q2H6-PQFN> : accessed 12 January 2024.

Images: Photograph. Port Arthur Convict Quarries. 12 October 2023. Port Arthur, Tasmania. Jeffrey Bollard, photographer. 2023. a. Private collection of author, Canberra.

Images: Photograph. Convict Memorial Tower. 6 October 2023. Hobart Penitentiary Chapel, Hobart, Tasmania. Jeffrey Bollard, photographer. 2023. b. Private collection of author, Canberra.

Obituaries. (1938) *GoulburnValley Stock and Proprty Journal*. 13 July. Victoria WOODS, J.T. p. 8. <https://trove.nla.gov.au/newspaper/article/219771008> : accessed 7 January 2024.

Testamentary records. England. 18 March 1720. ARNOLD, Coe. Will. Norwich Consistory Court Probate Records: Will Registers. NCC will register Archer 1. Norfolk Record Office, Norfolk, England. <https://nrocatalogue.norfolk.gov.uk/arnold-coe-beer-brewer-of-lowestoft-suffolk> : accessed 29 January 2024.

Unknown photographer. Images: Photograph. Image of the card catalogue at the National Library of Australia, showing the transition from the card catalogue to computer/microfiche. August 1980. Canberra. Copy of photograph provided to author by the National Library of Australia April 2024.

## Secondary Sources

- Abdelhady, Ahmad M. & Davis, Christopher R. (2023) Plastic Surgery and Artificial Intelligence: How ChatGPT Improved Operation Note Accuracy, Time, and Education. *Mayo Clinic Proceedings: Digital Health*. 1 (3). pp. 299–308.  
<https://www.sciencedirect.com/science/article/pii/S2949761223000470> : accessed 6 January 2024.
- Adrion, Jeffrey, Lang, Jenna, Noto, Keith, et al. (2023) *Ethnicity Estimate 2023 White Paper*. Ancestry.com. <https://www.ancestrycdn.com/support/us/2023/09/Ethnicity2023whitepaper.pdf> : accessed 5 February 2024.
- Ahnert, Ruth. (2023) *The Living with Machines Report [pdf version]*. Arts and Humanities Research Council. [https://livingwithmachines.ac.uk/wp-content/uploads/2023/11/AT1007%E2%80%94Living-With-Machines\\_FINAL-Fixed.pdf](https://livingwithmachines.ac.uk/wp-content/uploads/2023/11/AT1007%E2%80%94Living-With-Machines_FINAL-Fixed.pdf) : accessed 30 December 2023.
- Alharbi, Wardah S. & Rashid, Mamoon. (2022) A review of deep learning applications in human genomics using next-generation sequencing data. *Human Genomics*. 16 (1). pp. 5–20.  
<https://doi.org/10.1186/s40246-022-00396-x> : accessed 30 December 2023.
- Ancestry Hints®. [https://support.ancestry.com/s/article/Ancestry-Hints?language=en\\_US](https://support.ancestry.com/s/article/Ancestry-Hints?language=en_US) : accessed 4 February 2024.a.
- Ancestry Team. (2019) How Ancestry® Is Preserving History One Record at a Time. *Ancestry Blog*. [blog] 26 August. <https://blogs.ancestry.com/cm/ancestry-imaging-center-what-does-it-take-to-create-an-image-on-ancestry/> : accessed 10 January 2024.
- Ancestry Team. *Ancestry® to Apply Handwriting Recognition Artificial Intelligence to Create a Searchable Index of the 1950 U.S. Census*. <https://ancestry-prod2.adobecqms.net/content/ancestry-corp/en-us/blog/ancestry-apply-handwriting-recognition-artificial-intelligence-create-searchable-index-1950-us> : accessed 31 December 2023.
- Ancestry Team. *How SideView™ Technology Splits Your DNA Results by Parent*. <https://support.ancestry.com.au/s/article/SideView-Technology> : accessed 28 January 2024.
- AncestryDNA®. *Genotype*. <https://www.ancestry.com/c/dna-learning-hub/genotype> : accessed 28 January 2024.
- Australasian Association of Genealogists and Record Agents. *AAGRA Code of Ethics*. <https://www.aagra.asn.au/aagra-code-of-ethics/> : accessed 2 February 2024.
- Ball, Catherine A, Barber, Mathew J, Byrnes, Jake, et al. (2016) *AncestryDNA Matching White Paper: Discovering genetic matches across a massive, expanding genetic database*. Ancestry.com. <https://www.ancestry.com/cs/dna-help/matches/whitepaper> : accessed 4 February 2024.
- Bell, Genevieve, Zafiroglu, Alex, Assaad, Zena, et al. (2021) *Custodians & Midwives: The Library of the Future*. Canberra: Australian National University.  
[https://cybernetics.anu.edu.au/assets/Custodians\\_and\\_Midwives\\_FINAL\\_web.pdf](https://cybernetics.anu.edu.au/assets/Custodians_and_Midwives_FINAL_web.pdf) : accessed 3 October 2023.
- Benn, Ashley. (2023) *I am one of the many who is conflicted on whether or not to use AI to illustrate our genealogy*, Genealogy and Artificial Intelligence, [Facebook post]. 19 November.  
<https://www.facebook.com/groups/genealogyandai/permalink/1414556692487018> : accessed 26 January 2024.
- Bettinger, Blaine T. (2016) *The Family Tree Guide to DNA Testing and Genetic Genealogy: How to Harness the Power of DNA to Advance Your Family Tree Research*. London: Penguin.

- Bettinger, Blaine T. (2023a) Leveraging Artificial Intelligence Tools for Genealogical Research. *Association of Professional Genealogists Quarterly*. December 2023 pp. 8–13.
- Bettinger, Blaine T. (2023b) *This group is at the forefront of genealogy and AI*, Genealogy and Artificial Intelligence, [Facebook post]. 19 November. <https://www.facebook.com/groups/genealogyandai/permalink/1414732652469422> : accessed 26 January 2024.
- Board for Certification of Genealogists. *Genealogist's Code of Ethics*. <https://bcgcertification.org/ethics-standards/code/> : accessed 2 February 2024.
- Cain, David. (2023) Decoding the Human Blueprint: AI Revolutionizes Genetic Research. *Linked In*. <https://www.linkedin.com/pulse/genes-meet-machines-love-story-between-ai-genetic-research-david-cain/> : accessed 29 January 2024.
- Chatila, Raja, Firth-Butterfield, Kay, & Havens, John C. (2018) *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems Version 2*. Los Angeles: University of Southern California. <https://apps.dtic.mil/sti/citations/AD1170922> : accessed 1 February 2024.
- Christofides, Emily & O'Doherty, Kieran. (2016) Company disclosure and consumer perceptions of the privacy implications of direct-to-consumer genetic testing. *New Genetics and Society*. 35 (2). pp. 101–123. <https://www.tandfonline.com/doi/full/10.1080/14636778.2016.1162092> : accessed 29 January 2024.
- Cooke, Lisa. (2023) Using ChatGPT or Bard for Genealogy. *Genealogy Gems*. [blog] 22 June. <https://lialouisecooke.com/2023/06/22/ai-chatbots-genealogy/> : accessed 21 May 2024.
- Cowgill, Bo, Dell'Acqua, Fabrizio, Deng, Samuel, et al. (2020) 'Biased Programmers? Or Biased Data? A Field Experiment in Operationalizing AI Ethics.' In: *Proceedings of the 21st ACM Conference on Economics and Computation*, 21st Conference on Economics and Computation. New York, NY, USA: Association for Computing Machinery. pp. 679–681. <https://dl.acm.org/doi/10.1145/3391403.3399545> : accessed 4 February 2024.
- Cowley, Trudy, Frost, Lucy, Inwood, Kris, et al. (2021) Reconstructing a Longitudinal Dataset for Tasmania. *Historical Life Course Studies*. 11 pp. 20–47. <https://hlcs.nl/article/view/10912> : accessed 2 March 2024.
- Creswell, John W. & Creswell, J. David. (2022) *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. [Kindle Version]. 6th ed. edition. Los Angeles: SAGE Publications. <https://www.amazon.ca> : accessed 15 January 2024.
- Crymble, Adam. (2021) *Technology and the Historian: Transformations in the Digital Age*. [Kindle version]. Illinois: University of Illinois Press. [www.amazon.com.au](http://www.amazon.com.au) : accessed 17 December 2023.
- Dell'Acqua, Fabrizio. (2022) *Falling Asleep at the Wheel: Human/AI Collaboration in a Field Experiment on HR Recruiters: Working Paper*. USA: Laboratory For Innovation Science, Harvard Business School.
- Dell'Acqua, Fabrizio, McFowland, Edward, Mollick, Ethan, et al. (2023) Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality. *Harvard Business School Technology & Operations Mgt. Unit Working Paper*. No. 24-013 pp. 1–58. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4573321](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4573321) : accessed 7 January 2024.
- Denscombe, Martin. (2021) *The Good Research Guide: Research methods for small-scale social research projects*. 7th ed. New York: Open University Press.
- Dias, Raquel & Torkamani, Ali. (2019) Artificial intelligence in clinical and genomic diagnostics. *Genome Medicine*. 11 (1). p. 70. <https://doi.org/10.1186/s13073-019-0689-8> : accessed 8 December 2023.

- Duff, Wendy M. & Johnson, Catherine A. (2003) Where Is the List with All the Names? Information-Seeking Behavior of Genealogists. *The American Archivist*. 66 (1). pp. 79–95. <https://www.jstor.org/stable/40294218> : accessed 7 January 2024.
- Duff, Wendy M. & Johnson, Catherine A. (2002) Accidentally Found on Purpose: Information-Seeking Behavior of Historians in Archives. *The Library Quarterly: Information, Community, Policy*. 72 (4). pp. 472–496. <https://www.jstor.org/stable/40039793> : accessed 7 January 2024.
- Edwards, Lilian. (2022) *The EU AI Act proposal: a summary of its significance and scope*. London, UK: Ada Lovelace Institute. <https://www.adalovelaceinstitute.org/resource/eu-ai-act-explainer/>.
- Erica. *Introducing AI Record Finder™, the World's First AI Chat-Based Search Engine for Historical Records*. <https://blog.myheritage.com/2023/12/introducing-ai-record-finder-the-worlds-first-ai-chat-based-search-engine-for-historical-records/> : accessed 31 December 2023.
- Facebook. *Genealogy and Artificial Intelligence (AI) | Facebook Group*. <https://www.facebook.com/groups/genealogyandai> : accessed 23 April 2024.
- FamilySearch. *Wiki: Netherlands Civil Registration Records*. [https://www.familysearch.org/en/wiki/Netherlands\\_Civil\\_Registration\\_Record](https://www.familysearch.org/en/wiki/Netherlands_Civil_Registration_Record) : accessed 27 May 2024.
- FamilySearch. *RootsTech Videos 2023-2024*. <https://www.familysearch.org/en/rootstech/> : accessed 21 April 2024.
- FamilySearch. *Full-Text Search*. <https://www.familysearch.org/search/full-text> : accessed 10 April 2024.
- Fjeld, Jessica, Achten, Nele, Hilligoss, Hannah, et al. (2020) *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI*. Rochester, NY: The Berkman Klein Center for Internet & Society Research. <https://papers.ssrn.com/abstract=3518482> : accessed 31 January 2024.
- Gabriel, John. (2023) *Has anyone tried ChatGPT for genealogy yet? - WikiTree G2G*. WikiTree G2G, <https://www.wikitree.com/g2g/1550687/has-anyone-tried-chatgpt-for-genealogy-yet> : accessed 22 May 2024.
- Genealogy and Artificial Intelligence Facebook Group*. <https://www.facebook.com/groups/1255245945084761> : accessed 20 October 2023.b.
- Genetic Genealogy Standards Committee. *Genetic Genealogy Standards*. <http://www.geneticgenealogystandards.com/> : accessed 5 April 2024.
- Gibbs, Frederick. (2016) New Forms of History: Critiquing Data and Its Representations. *The American Historian*. 7 pp. 31–36. <https://www.oah.org/tah/february-2/new-forms-of-history-critiquing-data-and-its-representations/> : accessed 7 March 2024.
- Gibbs, Martin D., Tuffin, Richard L., Maxwell-Stewart, Hamish, et al. (2018) Landscapes of production and punishment: convict labour management on the Tasman Peninsula 1830–1877. *Antiquity*. 92 (362). <https://www.cambridge.org/core/journals/antiquity/article/landscapes-of-production-and-punishment-convict-labour-management-on-the-tasman-peninsula-18301877/A33CBE5287D38F307BFC526871276BB3> : accessed 1 March 2024.
- Glynn, Claire L. (2022) Bridging Disciplines to Form a New One: The Emergence of Forensic Genetic Genealogy. *Genes*. 13 (8). pp. 1–16. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9407302/> : accessed 27 January 2024.
- Godfrey, Barry, Homer, Caroline, Inwood, Kris, et al. (2021) Crime, Penal Transportation, and Digital Methodologies. *Journal of World History*. 32 (2). University of Hawai'i Press. pp. 241–260. <https://muse.jhu.edu/pub/5/article/794329> : accessed 27 February 2024.

- Goodfellow, Ian, Bengio, Yoshua, & Courville, Aaron. (2016) *Deep Learning [pdf version]*. Cambridge, USA: The MIT Press. <https://www.deeplearningbook.org/> : accessed 23 January 2024.
- Hanacek, N. *How Do Companies Measure DNA to Discover Your Ancestry?*. <https://www.nist.gov/how-do-you-measure-it/how-do-companies-measure-dna-discover-your-ancestry> : accessed 31 December 2023.
- Heraldry and Genealogy Society of Canberra. (1988) *Family History Beginners and Beyond*. 4th ed. Canberra ACT: Elect Printing.
- Heraldry and Genealogy Society of Canberra. (2002) *Family History Beginners and Beyond*. 12th ed. Canberra ACT: Australian National University Printing Services.
- Heraldry and Genealogy Society of Canberra. (2020) *Family History Beginners and Beyond*. 16th ed. Canberra ACT: Australian National University Printing Services.
- Hoitink, Yvette. (2023) Five good and bad ways to use ChatGPT for Genealogy. *Dutch Genealogy*. [blog] 16 April. <https://www.dutchgenealogy.nl/five-good-and-bad-ways-to-use-chatgpt-for-genealogy/> : accessed 16 May 2024.
- Hoitink, Yvette. (2024) 'Handwritten Text Recognition: Harnessing the Power of Artificial Intelligence to Transcribe and Search.' In: *Session 203 Syllabus , Expanding Possibilities*. Virginia, USA: National Genealogical Society. <https://www.ngsgenealogy.org/> : accessed 18 May 2024.
- Holton, Graham S. (2019) *Tracing Your Ancestors Using DNA: A Guide for Family Historians.[ Kindle Version]*. Pen and Sword. <https://www.amazon.ca/> : accessed 3 February 2024.
- Holton, Graham S. (2020) Discovering unknown medieval descents : a genetic approach – medieval genealogy for the masses. *Foundation*. 12 pp. 2–15. [https://strathprints.strath.ac.uk/73099/1/Holton\\_Foundations\\_2020\\_Discovering\\_unknown\\_medieval\\_descents\\_a\\_genetic\\_approach\\_medieval\\_genealogy\\_for\\_the\\_masses.pdf](https://strathprints.strath.ac.uk/73099/1/Holton_Foundations_2020_Discovering_unknown_medieval_descents_a_genetic_approach_medieval_genealogy_for_the_masses.pdf) : accessed 4 February 2024.
- Hörnemalm, Adam. (2023) *ChatGPT as a Software Development Tool : The Future of Development*. Unpublished Master's Thesis, Umeå University. <https://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-209909> : accessed 22 October 2023.
- Hosseini, Mohammad, Holmes, Kristi, & Resnick, David. (2023) The ethics of disclosing the use of artificial intelligence tools in writing scholarly manuscripts - Mohammad Hosseini, David B Resnik, Kristi Holmes, 2023. *Research Ethics*. 19 (4). p. pp.449-465. <https://journals-sagepub-com.proxy.lib.strath.ac.uk/doi/10.1177/17470161231180449> : accessed 2 May 2024.
- Hu, Krystal. *ChatGPT sets record for fastest-growing user base - analyst note*. <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/> : accessed 21 May 2024.
- Hu, Rui, Maier, Paul, Runfeldt, Goran, et al. (2021) *Family Finder Matching 5.0: Matching Algorithm and Relationship Estimation White Paper*. FamilyTreeDNA. [https://blog.familytreedna.com/wp-content/uploads/2021/08/Family\\_Finder\\_Matching\\_WhitePaper.pdf](https://blog.familytreedna.com/wp-content/uploads/2021/08/Family_Finder_Matching_WhitePaper.pdf) : accessed 4 February 2024.
- ISOGG Wiki. *Ethics, Guidelines and Standards*. [https://isogg.org/wiki/Ethics,\\_guidelines\\_and\\_standards](https://isogg.org/wiki/Ethics,_guidelines_and_standards) : accessed 2 February 2024.
- Jiang, Shuo, Luo, Jianxi, Hu, Jie, et al. (2022) Deep Learning for Technical Document Classification. *arXiv*. <https://arxiv.org/ftp/arxiv/papers/2106/2106.14269.pdf>.
- Johnson, Elizabeth-Anne. (2017) *Toward a collaborative online framework for archival representation*. Unpublished Master's Thesis, University of Manitoba. [https://www.academia.edu/34550820/Toward\\_a\\_collaborative\\_online\\_framework\\_for\\_archival\\_representation](https://www.academia.edu/34550820/Toward_a_collaborative_online_framework_for_archival_representation) : accessed 28 December 2023.

- Kaebnick, Gregory E., Magnus, David Christopher, Kao, Audiey, et al. (2023) Editors' Statement on the Responsible Use of Generative AI Technologies in Scholarly Journal Publishing. *Hastings Center Report*. 53 (5). pp. 3–6. <https://onlinelibrary.wiley.com/doi/abs/10.1002/hast.1507> : accessed 2 May 2024.
- Kaminski, June. (2011) Diffusion of Innovation Theory. *Canadian Journal of Nursing Informatics*. 6 (2). <https://cjni.net/journal/?p=1444> : accessed 10 January 2024.
- Kellogg, Katherine, Lifshitz-Assaf, Hila, Randazzo, Steven, et al. (2024) Don't Expect Juniors to Teach Senior Professionals to Use Generative-AI: Emerging Technology Risks and Novice AI Risk Mitigation Tactics. *Harvard Business School Technology & Operations Mgt. Unit Working Paper*. 24–074 [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4857373](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4857373) : accessed 6 December 2024.
- Kranzberg, Melvin. (1986) Technology and History: "Kranzberg's Laws." *Technology and Culture*. 27 (3). pp. 544–560. <https://www.jstor.org/stable/3105385> : accessed 2 February 2024.
- Lee, Devon Noel. (2023) ChatGPT Failed? AI Struggles to Write Family History. *Family History Fanatics*. [blog] 12 June. <https://www.familyhistoryfanatics.com/chatgpt-writes-family-history> : accessed 22 May 2024.
- Leeds, Dana. (2023) Using Artificial Intelligence in Genealogy. *Dana Leeds*. [blog] 24 October. <https://www.danaleeds.com/aigenealogy/> : accessed 21 May 2024.
- Li, Fei-Fei. (2023) *The Worlds I See: Curiosity, Exploration, and Discovery at the Dawn of AI*. [Kindle Version]. New York: Flatiron Books. <http://amazon.com.au> : accessed 1 February 2024.
- Little, Steve. (2024a) The Next Frontier: AI and Large Language Models Transforming Genealogy. *NGS Magazine*. 50 (1). pp. 20–24.
- Little, Steve. *AI Genealogy Insights*. <https://aigenealogyinsights.com/author/digitalarchivist/> : accessed 4 February 2024.a.
- Little, Steve. (2023b) 'Empowering Genealogists with Artificial Intelligence: Syllabus Week 2.' National Genealogical Society (US) course. USA.
- Little, Steve. (2024b) Avoiding Extremism: The Use and Disclosure of AI in Genealogy. *AI Genealogy Insights*. [blog] 9 June. <https://aigenealogyinsights.com/2024/06/09/avoiding-extremism-the-use-and-disclosure-of-ai-in-genealogy/> : accessed 10 June 2024.
- Little, Steve. (2024c) *Congrats to Dana Stewart Leeds and Thomas MacEntee for Top 10 finishes with AI topics for March 2024 Legacy Family Tree Webinars* Genealogy and Artificial Intelligence (AI), [Facebook post]. 2 April. <https://www.facebook.com/groups/genealogyandai/permalink/1489472768328743> : accessed 21 April 2024.
- Little, Steve. (2023c) New Use Case: Handwriting Recognition to Structured Data. *AI Genealogy Insights*. [blog] 9 October. <https://aigenealogyinsights.com/2023/10/09/new-use-case-handwriting-recognition-to-structured-data/> : accessed 17 October 2023.
- Lucy, Catherine. (2016) Research Trends & Emerging Technologies for Genealogists. *School of Information Student Research Journal*. 5 (2). pp. 1–23. <https://scholarworks.sjsu.edu/ischoolsrj/vol5/iss2/4> : accessed 9 December 2023.
- MacEntee, Thomas. *A Genealogist's Guide to Artificial Intelligence*. <https://familytreemagazine.com/resources/software/ai-and-genealogy/> : accessed 28 January 2024.
- Manikandan, Bala. *Demystifying ChatGPT: A Deep Dive into Reinforcement Learning with Human Feedback*. <https://bmanikan.medium.com/demystifying-chatgpt-a-deep-dive-into-reinforcement-learning-with-human-feedback-1b695a770014> : accessed 3 February 2024.



- Marr, Bernard. *A Short History Of ChatGPT: How We Got To Where We Are Today*.  
<https://www.forbes.com/sites/bernardmarr/2023/05/19/a-short-history-of-chatgpt-how-we-got-to-where-we-are-today/?sh=4c7e7cfb674f> : accessed 21 May 2024.
- Maxwell-Stewart, Hamish. (2022) *Making Crime Pay*. Australian Research Council Application. Project ID: LP220100118. Copy provided by author to Cheryl Bollard. October 2023.
- Maxwell-Stewart, Hamish. (2023) 'A Tutorial on the Founders and Survivors Database : Applications for genealogists.' Convict Special Interest Group. Family History ACT. Canberra ACT : accessed 9 March 2024.
- Ménard, Timothé. (2022) Good quality practices for artificial intelligence in genetics. *European Journal of Human Genetics*. 30 (9). pp. 993–995. <https://www.nature.com/articles/s41431-022-01061-6> : accessed 8 December 2023.
- Milligan, Ian. (2022) 'The Transformation of Historical Research in the Digital Age. In: Woolfe, Daniel. ed.' *Elements in Historical Theory and Practice*. Cambridge, UK: Cambridge University Press. pp. 1–38.
- Mollick, Ethan. (2024a) *Co-Intelligence: Living and Working with AI. [Kindle version]*. New York: Portfolio. <http://www.amazon.ca> : accessed 4 March 2024.
- Mollick, Ethan. (2024b) Innovation through prompting. *One Useful Thing*. [blog] 22 April. <https://www.oneusefulthing.org/p/innovation-through-prompting> : accessed 23 April 2024.
- Mollick, Ethan. (2024c) What Apple's AI Tells Us: Experimental Models. *One Useful Thing*. [blog] 11 June. <https://www.oneusefulthing.org/p/what-apples-ai-tells-us-experimental> : accessed 11 June 2024.
- Mollick, Ethan. (2024d) Doing Stuff with AI: Opinionated Mid-year Edition. *One Useful Thing*. [blog] 6 June. <https://www.oneusefulthing.org/p/doing-stuff-with-ai-opinionated-midyear> : accessed 6 June 2024.
- Mollick, Ethan. (2024e) What Can be Done in 59 Seconds: An Opportunity (and a Crisis). *One Useful Thing*. [blog] 31 January. [https://www.oneusefulthing.org/p/what-can-be-done-in-59-seconds-an?utm\\_campaign=post&fbclid=IwAR1hZDJdaCDG0mH3FRmf8NqY4P\\_F4ijuCMc9mlaiAGdupSU70XRpyZuCJvc](https://www.oneusefulthing.org/p/what-can-be-done-in-59-seconds-an?utm_campaign=post&fbclid=IwAR1hZDJdaCDG0mH3FRmf8NqY4P_F4ijuCMc9mlaiAGdupSU70XRpyZuCJvc) : accessed 2 February 2024.
- My Heritage. *Autoclusters - How to Use Autoclusters on MyHeritage*.  
<https://education.myheritage.com/article/autoclusters-for-dna-matches/> : accessed 31 May 2024.
- MyHeritage. *AI & Genealogy: Harnessing The Power Of Artificial Intelligence For Family History Research - MyHeritage Knowledge Base*. <https://education.myheritage.com/article/ai-genealogy-harnessing-the-power-of-artificial-intelligence-for-family-history-research/> : accessed 22 May 2024.
- Open AI. *ChatGPT — Release Notes*. <https://help.openai.com/en/articles/6825453-chatgpt-release-notes> : accessed 11 June 2024.
- Open AI. (2023a) *GPT-4 System Card*. Open AI. <https://cdn.openai.com/papers/gpt-4-system-card.pdf> : accessed 3 February 2024.
- Open AI. *Introducing GPT-4o and more tools to ChatGPT free users*. <https://openai.com/index/gpt-4o-and-more-tools-to-chatgpt-free/> : accessed 27 May 2024.b.
- Open AI. (2023b) *GPT-4 Technical Report*. Open AI. <https://cdn.openai.com/papers/gpt-4.pdf> : accessed 3 February 2024.
- Open AI. *Research*. <https://openai.com/research/> : accessed 19 May 2024.
- Open Culture. (n.d.) The Venice Time Machine: 1,000 Years of Venice's History Gets Digitally Preserved with Artificial Intelligence and Big Data. *Open Culture*. [blog]  
<https://www.openculture.com/2019/03/the-venice-time-machine.html> : accessed 3 June 2024.

- Parikh, Ravi B., Teeple, Stephanie, & Navathe, Amol S. (2019) Addressing Bias in Artificial Intelligence in Health Care. *JAMA*. 322 (24). pp. 2377–2378. <https://doi.org/10.1001/jama.2019.18058> : accessed 29 January 2024.
- Pillai, Abhilash. (2023) The Intersection of AI, Graph Neural Networks, and Genealogy: A New Era of Discovery. *Linked In*. <https://www.linkedin.com/pulse/intersection-ai-graph-neural-networks-genealogy-new-era-pillai> : accessed 8 October 2023.
- Register of Qualified Genealogists. *Professional Code*. <https://www.qualifiedgenealogists.org/the-professional-code> : accessed 19 February 2024.
- Rogers, Everett M. (2003) *Diffusion of Innovations, 5th Edition*. 5th edition. New York: Free Press.
- Royal, Charmaine D., Novembre, John, Fullerton, Stephanie M., et al. (2010) Inferring Genetic Ancestry: Opportunities, Challenges, and Implications. *American Journal of Human Genetics*. 86 (5). pp. 661–673. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2869013/> : accessed 30 January 2024.
- Sherratt, Tim. (2019) 'Hacking Heritage: Understanding the limits of online access. In: Lewi, Hannah and Smith, Wally. eds.' *The Routledge International Handbook of New Digital Practices in Galleries, Libraries, Archives, Museums and Heritage Sites*. Routledge. <https://doi.org/10.4324/9780429506765> : accessed 8 March 2024.
- Society of Australian Genealogists. *Ironclad Sisterhood: Convict Women in Australia*. <https://ironcladsisterhood.sag.org.au/> : accessed 9 March 2024.
- Southard, Diahann. (2021) Understanding AncestryDNA Genetic Communities. *Your DNA Guide - Diahann Southard*. [blog] 1 February. <https://www.yourdnaguide.com/ydgblog/ancestrydna-genetic-communities> : accessed 31 May 2024.
- Stallard, Matthew & de Groot, Jerome. (2020) “Things Are Coming Out That Are Questionable, We Never Knew About”: DNA and the New Family History. *Journal of Family History*. 45 (3). pp. 274–294. <https://doi.org/10.1177/0363199020906853> : accessed 10 January 2024.
- Stankus, Tony. (2020) Reviews of Science for Science Librarians: Direct-to-Consumer DNA Testing for Ancestry as a Complement to Traditional Genealogical Methods. *Science & Technology Libraries*. 39 (3). pp. 227–243. <https://doi.org/10.1080/0194262X.2020.1758285> : accessed 27 January 2024.
- Survey Monkey. *Pearson vs Spearman correlations: practical applications*. <https://www.surveymonkey.com/market-research/resources/pearson-correlation-vs-spearman-correlation/> : accessed 6 April 2024.
- Tanner, James L. (2024) 'Using Artificial Intelligence Tools to Expand your Genealogical Research Universe.' RootsTech 2024. Salt Lake City, Utah, USA <https://www.familysearch.org/en/rootstech/session/using-artificial-intelligence-tools-to-expand-your-genealogical-research-universe-part-1> : accessed 9 March 2024.
- Tibbo, Helen. (2003) Primarily History in America: How U.S. Historians Search for Primary Materials at the Dawn of the Digital Age. *The American Archivist*. 66 (1). pp. 9–50. <https://doi.org/10.17723/aarc.66.1.b12037011g718n74> : accessed 27 December 2023.
- Tuffin, Richard. (2023) 'Using Digital Technologies to learn more about Convicts.' Convict Special Interest Group. Family History ACT. Canberra ACT : accessed 9 March 2024.
- Tuffin, Richard, Gibbs, Martin, Roberts, David, et al. (2018) Landscapes of Production and Punishment: Convict labour in the Australian context. *Journal of Social Archaeology*. 18 (1). pp. 50–76. <https://journals.sagepub.com/doi/epub/10.1177/1469605317748387> : accessed 21 October 2023.
- University of Exeter. *The Material Culture of Wills, England 1540-1790*. <https://sites.exeter.ac.uk/materialcultureofwills/about/> : accessed 9 April 2024.

- University of New England. *World-first digital project gives unique insights into convict life*. <https://www.une.edu.au/connect/news/2023/11/world-first-digital-project-gives-unique-insights-into-convict-life> : accessed 27 February 2024.
- Walsh, Toby. (2023) *Faking It: Artificial Intelligence in a Human World*. Melbourne, Victoria: La Trobe University Press.
- Whitford, Sara. (2023) WARNING! Be cautious using ChatGPT for genealogy research. *East Carolina Roots*. [blog] 27 March. <https://www.eastcarolinaroots.com/warning-be-cautious-using-chatgpt-for-genealogy-research/> : accessed 22 May 2024.
- Yakel, Elizabeth. (2004) Seeking Information, Seeking Connections, Seeking Meaning: Genealogists and Family Historians. *Information Research*. 10 (1). <https://informationr.net/ir/10-1/paper205.html> : accessed 4 July 2024.
- Yin, Robert K. (2003) *Case Study Research: Design and Methods*. 3rd ed. Cambridge, USA: SAGE Publications, Inc. [https://iwansuharyanto.files.wordpress.com/2013/04/robert\\_k-\\_yin\\_case\\_study\\_research\\_design\\_and\\_mebookfi-org.pdf](https://iwansuharyanto.files.wordpress.com/2013/04/robert_k-_yin_case_study_research_design_and_mebookfi-org.pdf).