

<https://doi.org/10.19164/jlrm.v5i1.1983>

First online date: 04/10/2026



Article

The Reflexive Local AI (RLAI) Framework: a methodology for AI-assisted transcription in empirical legal research

Lloyd Brown*

Swansea University

<https://orcid.org/0000-0003-0945-7472>

Ryan Brown**

Abstract

Automated transcription has been increasingly adopted to relieve the burden of manual transcription in qualitative research, including empirical legal scholarship. Yet cloud-based commercial transcription services raise serious data security and confidentiality concerns, particularly where terms of service reserve rights to use uploaded content for model training. Existing scholarship, notably Da Silva's work extending Bokhove and Downey, has identified these vulnerabilities and responded with a risk management framework grounded in security practice. This article departs from that framing. It grounds the case for local processing not merely in prudent risk mitigation, but in the specific legal architecture of UK

* Lecturer in Law, HRC School of Law, Swansea University. Email: lloyd.a.brown@swansea.ac.uk. ORCID: 0000-0003-0945-7472. Incoming Lecturer in Law at the School of Law and Politics, Cardiff University (December 2026). This research was generously supported by a British Academy / Leverhulme Small Research Grant under the British Accounting & Finance Association (BAFA) special fund, for the project 'Pensions in a Warming World: An Empirical Study of the Opportunities and Challenges' (grant awarded October 2025, ref number: SRG/250023). The authors would also like to thank Dr Paul Dargue and the anonymous reviewers for their helpful comments and suggestions on an earlier draft of this article.

** Senior Cloud Engineer.

data protection law, arguing that local processing of interview data may constitute a necessary technical and organisational measure under Article 32 UK GDPR, in the absence of a vetted, institutionally governed processing arrangement. Drawing on the author's experience conducting thirty-three elite interviews for a British Academy/Leverhulme-funded project on pension governance and climate risk, this article develops that legal analysis into the Reflexive Local AI (RLAI) framework: five interlocking principles addressing local processing, human interpretive oversight, contextual anonymisation, transparent documentation and reflexive engagement. The article situates this framework within the methodological literature on transcription as constructivist interpretation, before analysing its legal foundations under Articles 5, 28 and 32 UK GDPR and Chapter V's international transfer regime. It concludes by identifying the framework's practical and empirical limits, and by proposing directions for its extension beyond transcription into AI-assisted qualitative analysis.

Article Text

1. Introduction

In recent years, automated transcription has been more greatly explored to circumvent the 'tedious and time-consuming' nature of manual transcription.¹ While cloud-based transcription services are effective, they can nonetheless introduce serious data security and ethical challenges for empirical researchers.² Inspired by his operational risk management approach, this article nevertheless departs from Da Silva's security-practitioner framing.³ It

1 Joseph Da Silva, 'Producing "good enough" automated transcripts securely: Extending Bokhove and Downey (2018) to address security concerns' (2021) *Methodological Innovations* 1-11, 1-2.

2 *ibid* 1. See also Regulation (EU) 2016/679 (UK GDPR), rec 6, which states: 'Rapid technological developments and globalisation have brought new challenges for the protection of personal data. The scale of collection and sharing of personal data has increased significantly.'

3 *ibid*.

does this by grounding the case for local processing in the specific legal architecture of UK data protection law, and by developing that analysis into a transferable methodological framework.

This work draws upon my experience as Principal Investigator of a British Academy/Leverhulme-funded project examining fiduciary duties and climate change in UK pension governance. Faced with prohibitive commercial transcription costs and significant confidentiality obligations, the second author (a senior cloud engineer) and I designed an entirely local artificial intelligence (AI) methodology using offline speech recognition and human-led verification. The specific details of this project, and the operational pressures that motivated this workflow, and the second author's role and its boundaries, are set out in section 2 below.

The central argument advanced in this article is legal as much as methodological. I argue that local 'processing'⁴ of interview data is not merely a prudent risk mitigation choice but, in the absence of a vetted, institutionally governed processing arrangement, a necessary technical and organisational measure to satisfy 'security of processing' obligations under Article 32 of Regulation (EU) 2016/679 (UK GDPR).⁵ Specifically, I argue that consumer-facing commercial AI transcription services (whose terms of service routinely reserve rights to use uploaded

4 UK GDPR, art 4(2): 'processing' means 'any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction.'

5 UK GDPR, art 32.

content for model training) cannot straightforwardly satisfy the requisite appropriate safeguards the law requires.

Methodologically speaking, transcription is an interpretive rather than administrative act,⁶ and the safeguards required by both law and qualitative rigour converge. I have developed this convergence into an applied framework, titled the 'Reflexive Local AI' (RLAI) framework. This framework comprises of five principles for conducting localised AI-assisted transcription in empirical legal research in a lawful and defensible way.

To establish this working model, the article is structured as follows: section 2 below sets out the research project that sparked the RLAI framework; section 3 importantly examines why the use of AI changes qualitative research and data processing; section 4 analyses the data protection landscape under the existing legal regime; section 5 operationalises that legal analysis into the five core principles; and finally, section 6 evaluates the broader implications and future directions of the research. Technical hardware specifications and local execution steps are succinctly provided in appendix I.

2. The research project

Just before the start of the 25/26 academic year, I had an email from the British Academy congratulating me on being awarded a British Academy/Leverhulme Small Research Grant for

6 Elinor Ochs, 'Transcription as Theory' in Elinor Ochs and Bambi B Schieffelin (eds), *Developmental Pragmatics* (New York: Academic Press 1979) 167-168.

my project titled *Pensions in a Warming World: An Empirical Study of the Opportunities and Challenges*. I had been awarded near the full amount under the British Accounting & Finance Association's (BAFA) 'special fund'.⁷ This grant was designed to cover, *inter alia*, travel for in-person interviews with pension sector expert.

Having achieved ethical approval, I thereafter conducted thirty-three semi-structured 'elite' interviews between October 2025 and July 2026. The participants included a cross-section of institutional actors, including pension trustees, specialised pension lawyers, academic commentators, investment consultants and senior policy makers. I had hoped that all interviews would be held in-person, but it was sometimes necessary to hold interviews online because of participants' busy work schedules. I provided each participant with a standard information sheet and consent form and received their written consent.⁸ After the interview, interviewees were sent a debrief document, outlining what happens next.

McDowell's account of 'using connections' to secure access is instructive here. She highlights how fieldwork depends heavily on 'luck and chance, connections and networks, and the particular circumstances at the time.'⁹ Such spontaneity often yields unexpected results. While in London for my project, a chance conversation on the Underground with a fellow

⁷British Academy/Leverhulme Small Research Grants' (BA, no date)
<<https://www.thebritishacademy.ac.uk/funding/schemes/ba-leverhulme-small-research-grants/>> accessed 1 August 2026.

⁸ UK GDPR, art 4(11): 'consent' is defined as 'any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her.'

⁹ Linda McDowell, 'Elites in the city of London: some methodological considerations' (1998) 30(12) *Environment and Planning A* 2133-2146, 2135.

passenger (who conveniently worked in the pensions sector) after I took the wrong train led directly to new interview participants.

Beyond such chance encounters, however, most access was built more deliberately. When I began the project, I did not have any contacts in the field. LinkedIn was imperative for finding connections and building a list of potential interview participants. I signed up to LinkedIn Premium so that I could gain fifteen 'InMail' credits each month, allowing me to contact anyone via the social media's messaging system. As my sample 'snowballed',¹⁰ I found that I relied less on LinkedIn and more on email. I learned that, for targeting elite participants, LinkedIn is an important tool for gaining access, and I would recommend factoring the cost of the Premium account into any grant costings (so long as such a cost is permitted by the funder).

Separately, the idea to use a localised AI-assisted transcription software (in my case, OpenAI's Whisper) came when I was writing up my project proposal prior to submission to the British Academy. During the drafting of costings, I approached a transcription company and was quoted a price of approximately £6,000 for twenty interviews (or roughly £3.50 plus VAT per audio minute). Such a fee would have inevitably wiped out a large proportion of my proposed funding and prohibited the reach of my fieldwork. As a full-time academic, I was also sufficiently worried about the work involved with manual transcription; in fact, I felt as though Sisyphus had a less arduous task to contend with. When I discussed these concerns with the

10 P Biernacki and D Waldorf, 'Snowball Sampling – Problems and Techniques of Chain Referral Sampling' (1981) 10 Sociological Methods & Research 141-163.

second author, a senior cloud engineer, he suggested using a localised AI to assist transcription. I had never considered such an idea but, so long as confidentiality was assured, I was willing to test the capacity of the new technology. We ran some preliminary tests before accepting the idea, and these worked well. The second author's involvement was confined to this technical set-up stage: it occurred before the data were collected, and he has not been involved at any stage of data processing or seen any of the interview data collected in the study. When setting up Whisper, we used our own test samples of our own voices, which did not include the interview data. He appears as the co-author on this paper on the basis of his technical contribution and his input to some of the appendix I technical guidance. Once it was set up and tested on my machine, the use of Whisper was explained in my ethics application and throughout the project documentation.

There are two examples during my fieldwork where I discussed the methodology directly to interviewees, and each illustrates a different dimension of building participant trust. First, more than once, online participants asked (without any apparent awareness of the confidentiality implications) why we were not simply recording the meeting directly on the video conferencing software. This question was a reasonable one; video conferencing now offers built-in recording services that are often used in business meetings. However, I explained to participants that this would potentially be a breach of confidence, since the audio file and transcript would be stored on the video conference companies' cloud storage. One interviewee found it amusing that I was using my 'old fashioned', hand-held Sony recording device. I myself sensed the juxtaposition between using the traditional handheld recorder, while later deploying a GenAI model to transcribe and time-stamp the audio files (something

that would have been unimaginable even when I was doing my doctoral research some ten years ago). However, the handheld recording device allowed me to save the data locally.

Secondly, I specifically recall a participant asking me about the transcription process itself. I explained that the AI software would run entirely locally, which was a technical detail I found meaningful in building trust. Several participants (this one included) seemed reassured once told that their testimony would remain on my own hardware rather than passing through commercial servers. This encounter is a small but concrete instance of the ‘positional strategy’ identified below.

3. Why AI changes qualitative methodology and vulnerabilities

This section establishes the methodological foundations upon which the article is built.

3.1. Elite interviews and transcription as constructivist interpretation

The literature surrounding ‘elite interviews’ has grown over the last three decades.¹¹ Empirical legal research frequently relies on elite semi-structured interviews. Harvey notes that there ‘is no clear-cut definition for the term “elite” and given its broad understanding across the social sciences, scholars have tended to adopt different approaches.’¹² Similar to my study, merchant and investment bankers based in the City of London, have been described as an ‘elite’ interview group.¹³ While the definition is elusive, I define ‘elite’ participants in this study by their institutional authority, specialist expertise or influence within pension governance. In

11 William S Harvey, ‘Strategies for conducting elite interviews’ (2011) 11(4) Qualitative Research 431-441, 432.

12 *ibid.*

13 McDowell (n 9).

particular, their professional visibility and membership of a relatively close-knit community create confidentiality challenges, as contextual details may make them identifiable even where names and institutional affiliations are removed from the transcripts.

A transcript does not capture objective reality but is a selective representation of an acoustic event.¹⁴ Ochs showed the subjectivity involved in transcription when she famously observed in 1979: 'A transcript that is too detailed is difficult to follow and assess. A more useful transcript is a more selective one.'¹⁵ Thus transcription is 'theory-laden'.¹⁶ The choices researchers make in producing the transcript enact the theoretical assumptions they hold and constrain the interpretations that can subsequently be drawn from the data.¹⁷

When speech recognition is delegated to AI, the researcher must remember that the need to perform interpretive construction does not disappear. Modern ASR systems do not 'copy' audio signals; instead, they process raw acoustic data through probabilistic language models that predict and output the most statistically likely sequence of words based on their training corpora.¹⁸ As Bender et al. highlight in their critique of natural language processing (NLP), large language models (LLMs) operate as 'stochastic parrots'.¹⁹ They join together plausible-

14 Ochs (n 6) 168.

15 *ibid.*

16 Judith C Lapadat and Anne C Lindsay, 'Transcription in Research and Practice: From Standardization of Technique to Interpretive Positionings' (1999) 5(1) *Qualitative Inquiry* 64-86, 77.

17 *ibid* 66.

18 Alec Radford, Jong Wook Kim, Tao Xu, Greg Brockman, Christine McLeavey and Ilya Sutskever, 'Robust Speech Recognition via Large-Scale Weak Supervision' (2022) arXiv <<https://cdn.openai.com/papers/whisper.pdf>> accessed 1 August 2026.

19 Emily M Bender, Timnit Gebru, Angelina McMillan-Major and Shmargaret Shmitchell, 'On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?' (2021) *Proc FAccT* 610, 616.

sounding sequences based on statistical co-occurrence without any underlying communicative intent, semantic understanding or contextual awareness.²⁰ This is likely compounded by the so-called ‘black box problem’.²¹

Researchers must increasingly be open to the possibility that AI may lead to the emergence of ‘ethical risk’ in qualitative research.²² Parasuraman and Manzey’s foundational work on automation complacency shows that automation does not simply replace human judgment but reshapes it, often in ways the designers never anticipated.²³ Both over-reliance and unwarranted disregard for automated output are common failures.²⁴ It is because of the risk of automation bias and misuse/disuse that I wished to keep a ‘human-in-the-loop’ during data verification. However, I entirely agree with Moore, who suggests that automated transcription technology, while lacking certain capabilities and liable to being inaccurate at times, may be highly beneficial for ‘providing first-pass transcripts.’²⁵

3.2. Expanding upon Da Silva

Bokhove and Downey recognised the challenges researchers face when utilising third-party automated captioning or transcription tools; but their paper did not prescribe a procedural

20 *ibid* 617.

21 Kyle ML Jones, ‘Generative AI in Qualitative research and Related Transparency Problems: A Novel Heuristic for Disclosing Uses of AI’ (2025) 24 *International Journal of Qualitative Methods* 1-14. Matthew Kosinski, ‘What is black box artificial intelligence (AI)?’ (IBM, 29 October 2024) <<https://www.ibm.com/think/topics/black-box-ai>> accessed 1 August 2026.

22 Kyungmee Lee, ‘Becoming Ethical Qualitative Researchers in the Era of Artificial Intelligence: Michel Foucault’s Ethics of the Self’ (2026) 32(2) *Qualitative Inquiry* (forthcoming/online first).

23 Raja Parasuraman and Dietrich H Manzey, ‘Complacency and Bias in Human Use of Automation: An Attentional Integration’ (2010) 53(3) *Human Factors: The Journal of the Human Factors and Ergonomics Society* 381-410.

24 *ibid* 381.

25 Robert J Moore, ‘Automated Transcription and Conversation Analysis’ (2015) 48(3) *Research on Language and Social Interaction* 253-270, 253.

‘solution’ beyond acquiring ethics committee approval.²⁶ Da Silva has drawn attention to the severe data security and confidentiality vulnerabilities introduced by commercial cloud-based transcription services.²⁷ He observes that even a cursory review of the terms and conditions of common zero-cost cloud transcription services shows they cannot offer adequate assurances of confidentiality for interview data promised anonymity by the researcher.²⁸

Attempting to bypass this cloud vulnerability through local device processing, Da Silva notes the severe practical barriers of proprietary hardware requirements, such as purchasing a dedicated Pixel device to run local recurrent neural networks, which still fail to offer a scalable methodological remedy for researchers.²⁹ On Bokhove and Downey’s ethics committee solution, Da Silva goes further, concluding that their specific method would be unusable in a considerable number of cases involving genuinely confidential data.³⁰ His response to these vulnerabilities, however, is grounded in risk management rather than statutory obligation; he does not ask whether, or under what conditions, local processing might be legally required rather than merely prudent. In this article, I address that critical gap.

3.3. Vulnerabilities in the study

While the use of AI presents methodological risks across all qualitative disciplines, empirical legal scholarship is perhaps uniquely exposed. During my research study I identified three specific vulnerabilities, as follows:

26 Christian Bokhove and Christopher Downey, ‘Automated generation of “good enough” transcripts as a first step to transcription of audio-recorded data’ (2018) *Methodological Innovations* 1-14, 12.

27 Da Silva (n 1).

28 *ibid* 2.

29 *ibid* 3.

30 *ibid*.

Vulnerability	Description
Specialised socio-legal terminology	Legal and financial discourse relies heavily on dense jargon, statutory references and acronyms (e.g. ‘TCFD’, ‘PSA26’, ‘fiduciary standards’). While overall accuracy was high,³¹ I found that the AI repeatedly misrecognised central domain terms such as ‘fiduciary’ (frequently rendering this word as ‘fed-user’, ‘federal’, ‘judiciary’, ‘FIDU3Duty’, and ‘5-2-3-2-2’), illustrating the necessity of domain-specific human verification. I found that the accuracy of the AI was dependent on the level of audible background noise and participants’ accents and speech patterns. This is something that Da Silva also found; he states that ‘Accuracy appeared to be reduced when transcribing a Scottish accent and improved when transcribing an American one.’³² In my research, for instance, Whisper had to be set with an ‘English’ parameter – early use led to the audio being transcribed entirely in Welsh, because of my Welsh accent.

³¹ It is noteworthy that most of my interview recordings were, on average, one hour long and the transcribed texts often spanned 10,000+ words. The longest interview transcribed was 91 minutes long and constituted 16,000+ words, or 48 pages. Overall, I was impressed by the AI’s speed in transcribing the audio: as a rough estimate, one minute of audio would be transcribed in one minute by the model. Despite the size of these transcripts and the speed in which it performed its work, I would only identify one or two minor mistakes in the AI’s transcription per recording. One recording was a little more challenging as the research interview took place in an area with sufficient background noise; but even so, the AI model still performed well.

³² Da Silva (n 1) 7. Da Silva usefully shows (in Tables 4 and 5) the differences in time taken to transcribe five interviews using a fully manual method vis-à-vis a semi-automated method. The semi-automated method took hours less to transcribe audio with similar recording lengths. I cannot provide such a comparison, since I only used a semi-automated method in my study. From looking at Table 5 in Da Silva’s paper (the one showing the semi-automated method), the transcription times seem similar to the time taken in my study.

Subtle contextual identifiers	Elite interviewing testimonies often contain indirect or contextual identifiers. A participant may not explicitly state their name or institution, but may reference specific regulatory filings, unique fund sizes or localised corporate disputes. The RLAI framework operates on the basis that AI-assistance should be limited to a first-pass transcription only, with further scrutiny being undertaken by the researcher. Unique to my study, but nevertheless a good example, I had particular issues when interviewees mentioned their Task Force on Climate-related Financial Disclosures (TCFD) reports,³³ which often contained fund-specific details, e.g., ‘Our approach changed significantly after our 2024 TCFD report, where we explicitly tied our net zero commitment to a phased divestment timeline, though we still hold a legacy equity stake in South Wales Oil Co. because of the income yield.’³⁴ To many in the elite group, the information in this statement is potentially revealing. This risk is not unique to legal research: McDowell, reflecting on her own interviews with City of London bankers, likewise concluded that her attempts at anonymisation might prove insufficient to disguise her respondents within banking’s ‘relatively small and close-knit world’.³⁵
-------------------------------	--

33 ‘Task Force on Climate-related Financial Disclosures’ (TCFD, no date) <<https://www.fsb-tcfd.org>> accessed 1 August 2026.

34 This is a fictitious example and is not linked to any person/institution.

35 McDowell (n 9) 2135-2137.

Power dynamics and trust	Conducting research with powerful institutional actors requires navigating complex access dynamics, ³⁶ and the trust built through that access must be sustained throughout the AI-assisted workflow. Assuring participants that their testimonies will be processed securely (example in section 2), without exposure to commercial cloud infrastructure, operates as a crucial positional strategy that preserves trust and protects participant confidentiality.
--------------------------	--

Table 1 - study vulnerabilities

Treating an AI-generated transcript as a transparent, objective representation of an elite legal interview ignores the theoretical and operational realities. AI outputs must instead be understood as a highly structured, first pass textual rendering that demands rigorous human audit, contextual anonymisation and reflexive interrogation.

4. The governance challenge

The legal governance of AI-assisted transcription begins with data protection law.³⁷ Because interview recordings constitute personal data (and may also reveal ‘special category’ data) researchers must ensure that every stage of AI processing complies with the UK GDPR³⁸ and the DPA 2018.³⁹ These remain the core statutory framework,⁴⁰ though the Data (Use and

36 Susan A Ostrander, ‘Surely You’re Not in This Just to Be Helpful: Access, Rapport, and Interviews in Three Studies of Elites’ (1993) 22(1) *Journal of Contemporary Ethnography* 7.

37 ‘Data protection’ (gov.uk, no date) <<https://www.gov.uk/data-protection>> accessed 1 August 2026.

38 *ibid.*

39 *ibid.*

40 Information Commissioner’s Office, ‘The Data Use and Access Act 2025 (DUAA) – what does it mean for organisations?’ (ico, 19 June 2026) <<https://ico.org.uk/about-the-ico/what-we-do/legislation-we-cover/data->

Access) Act 2025 (hereafter 'DUAA 2025') received Royal Assent on 19 June 2025, and its data protection provisions are now in force;⁴¹ where relevant, this article notes below whether specific provisions discussed have been affected.

4.1. Data protection as a right

The UK GDPR recognises the protection of personal data as a fundamental right,⁴² reflecting the continuing influence of Article 8 of the Charter of Fundamental Rights of the European Union.⁴³ That right is not absolute but must be balanced against other legitimate interests through the principle of proportionality.⁴⁴ For empirical researchers, this means that AI-assisted processing must be designed from the outset to comply with the data protection framework rather than treated as a purely technical or administrative matter. It is also worth noting, given the elite and institutional character of the sample discussed in section 2, that data protection compliance does not exhaust the relevant confidentiality analysis. Participants such as FCA-regulated professionals or pension trustees may owe, or be owed, sectoral or fiduciary confidentiality obligations that operate independently of, and in addition to, the UK GDPR.

[use-and-access-act-2025/the-data-use-and-access-act-2025-what-does-it-mean-for-organisations/](#)> accessed 1 August 2026.

41 Data (Use and Access) Act 2025 ('DUAA 2025').

42 UK GDPR, rec (1).

43 *ibid*, rec (1), citing Charter of Fundamental Rights of the European Union, art 8(1) and the Treaty on the Functioning of the European Union (TFEU), art 16(1).

44 *ibid*, rec (4).

4.2. The cloud privacy trap under the data protection regime

This analysis is directed at AI processing specifically, rather than institutional data storage as such. For my study, the audio files and transcripts were held on my university-managed cloud storage ('OneDrive'), which operates under Swansea University's own data processing agreement with Microsoft. This is a materially different arrangement from the consumer-facing AI platforms, since it carries no equivalent right to use file content for model training. And it is precisely this distinction, rather than cloud processing *per se* that the remainder of this section develops. A vetted, institutionally governed cloud arrangement of this kind is not what the necessity argument advanced here targets; it is, if anything, the confirming case, since the absence of a training data clause and the presence of proper institutional oversight are exactly what an unvetted commercial AI platform lacks. Indeed, reliance on consumer-facing commercial cloud AI platforms may create significant compliance risks under the UK GDPR. When a researcher uploads raw interview audio to these services, the data moves off local machines and onto servers controlled by commercial vendors. This transfer creates a legal conflict.

Under Article 4 of the UK GDPR, the employing university acts as the 'data controller', while the cloud vendor operates as a 'data processor'.⁴⁵ The definition of 'processor' is as follows: 'a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller.'⁴⁶ The UK GDPR further defines 'controller' as 'the natural or

⁴⁵ *ibid*, arts 4(7)-(8).

⁴⁶ *ibid*, 4(8).

legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data.’⁴⁷

The ‘controller’ and ‘processor’ classification is not fixed, however. A vendor only remains a processor for as long as it processes data strictly on the controller’s instructions. Where a vendor’s terms of service require the right to use uploaded content to train or refine its own models, it is processing that data for a purpose of its own (model improvement) rather than solely on the researcher’s instructions. This most straightforwardly constitutes a breach of the Article 28 processor relationship, and, depending on the facts, may on ordinary principles of controllership also risk tipping the vendor from processor into a controller in its own right for that purpose. Either characterisation leaves the researcher exposed: an Article 28 breach they cannot readily detect or remedy, or a compliance problem more serious than a processor merely falling short of its own obligations. The critical point for present purposes is not which of these two characterisations is correct, but that consumer-facing platforms of this kind cannot offer the researcher confidence that neither has occurred.

Moreover, Article 28(3)(a)-(h) requires binding contractual terms or other legal act under domestic law that limit processing strictly to the researcher’s instructions.⁴⁸ In reality, consumer-facing or low-cost AI platforms routinely include terms of service reserving rights to use uploaded content to train and refine their proprietary models.⁴⁹ Uploading unredacted

⁴⁷ *ibid*, 4(7).

⁴⁸ *ibid*, art 28(3)(a).

⁴⁹ Da Silva (n 1) 2.

interview files to train commercial models may be incompatible with Article 5(1)(b)'s 'purpose limitation' principle.⁵⁰ This principle states that personal data shall be 'collected (whether from the data subject or otherwise) for specified, explicit and legitimate purposes and not further processed by or on behalf of a controller in a manner that is incompatible with the purposes for which the controller collected the data ('purpose limitation').'⁵¹

Even where the relationship remains as a straightforward processor arrangement, a standard commercial terms of service rarely satisfies Article 28(3)'s substantive requirements. A terms of service is itself a contract,⁵² but Article 28(3) requires that contract to bind the processor to specific obligations.⁵³ For instance, Article 28(3)(a) states that the processor should only process 'the personal data on documented instructions from the controller'.⁵⁴ Such a transfer under the terms of service must be read with Article 32 in mind.⁵⁵ This deals with 'security of processing' and requires the controller and processor to implement 'appropriate technical and organisational measures to ensure a level of security appropriate to the risk'.⁵⁶ Consumer-facing and low cost AI platforms' terms of service are drafted to protect the vendor's commercial interests,⁵⁷ not to discharge GDPR obligations.

50 *ibid*, art 5(1)(b).

51 *ibid*.

52 Da Silva (n 1) 2. Here, Da Silva states: 'It was apparent from even a cursory reading of the terms and conditions of the predominant cloud-computing transcription offerings, for example Google... that these zero-cost offerings could not provide adequate assurances around the confidentiality of the data I was seeking to transcribe, and would not satisfy the ethical requirements of the research that I was performing, the anonymity of which I had agreed upfront with the participants.'

53 UK GDPR, art 28(3).

54 *ibid*, art 28(3)(a).

55 *ibid*, art 32.

56 *ibid*, art 32(1).

57 Da Silva (n 1) 2.

There is also the problem of cross-border data routing. The Information Commissioner's Office (ICO) states that while UK GDPR applies to most organisations in the UK, it is important and necessary to govern international transfers.⁵⁸ The ICO states: 'People risk losing the protection of UK data protection law if their personal information is transferred outside the UK.'⁵⁹ Under the UK GDPR, routing audio files to servers outside the UK or European Economic Areas (e.g. the United States) typically triggers 'Chapter V' restrictions on international transfers, since a transfer to any third country (not only an international organisation)⁶⁰ engages these provisions.⁶¹ The new rules relating to data transfers have been described as 'a key milestone in the DUAA's implementation'.⁶² Even following the amendments introduced by the DUAA 2025,⁶³ relying on cloud-based AI infrastructure still requires navigating complex transfer risk assessments and contractual safeguards.⁶⁴ Thus the need for localisation is not materially altered.

58 Information Commissioner's Office, 'What is an international transfer of personal information?' (ico, no date) <<https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/international-transfers/a-guide-to-international-transfers/what-is-an-international-transfer-of-personal-information/>> accessed 1 August 2026.

59 *ibid*

60 UK GDPR, art 4(26) defines an 'international organisation' as an organisation governed by public international law or established by an agreement between two or more countries. AI companies do not fall within this definition. However, Chapter V also applies to transfers to a 'third country', which encompasses transfers to overseas commercial providers.

61 *ibid*, art 44A(1)(a)-(b).

62 Amanda Leiu, 'The Data (Use and Access) Act 2025: Key Implications for Businesses as the New International Transfer Regime Takes Effect (Part 1)' (Burgess Salmon, 6 February 2026) <<https://www.burgess-salmon.com/articles/102mh1d/the-data-use-and-access-act-2025-key-implications-for-businesses-as-the-new-in/>> accessed 1 August 2026. See also Information Commissioner's Office, 'Are we making a restricted transfer' (ico, no date) <<https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/international-transfers/a-guide-to-international-transfers/are-we-making-a-restricted-transfer/>> accessed 1 August 2026.

63 DUAA 2025, s 85. See also 'Data (Use and Access) Act 2025: data protection and privacy changes' (gov.uk, 27 June 2025) <<https://www.gov.uk/guidance/data-use-and-access-act-2025-data-protection-and-privacy-changes>> accessed 1 August 2026; Alan Baker, 'Data (Use and Access) Act 2025: five key changes for businesses' (Farrer & Co, 5 May 2026) <<https://www.farrer.co.uk/news-and-insights/data-use-and-access-act-2025-five-key-changes-for-businesses/>> accessed 1 August 2026.

64 Alasdair Dougan and Sophie Burton-Jones, 'What are the new rules on international data transfers?' (Mills & Reeve, 5 February 2026) <<https://www.mills-reeve.com/publications/what-are-the-new-rules-on-international-data-transfers/>> accessed 1 August 2026.

4.3. Local deployment as a compliance-by-design response

The compliance hazards identified above are not inherent to AI-assisted transcription as such; they are a function of where processing occurs. Running speech recognition entirely offline, on hardware under the researcher's physical control, resolves each of the three problems identified above at source.⁶⁵ There is no transfer to a third-party processor, so no route by which data could be absorbed into a training corpus, and concomitantly, no engagement of the Chapter V international transfer regime.⁶⁶

This does not imply that every institutional cloud arrangement is equally deficient. Where a university holds an enterprise agreement with a commercial AI vendor that expressly excludes training data use and satisfies Article 28(3) in full, the analysis above would not apply in the same way. The point is narrower: consumer-facing platforms of the kind typically available to researchers without institutional procurement support cannot offer this assurance, and it is this class of arrangement, not cloud infrastructure as such, that the RLAI framework's first principle (localisation) is designed to guard against.

65 Nothing in this article should be taken as suggesting that the use of local AI transcription either automatically requires or dispenses with the need for a Data Protection Impact Assessment (DPIA). Under Article 35 UK GDPR, the requirement for a DPIA is determined by whether the proposed processing is 'likely to result in a high risk to the rights and freedoms of natural persons.' Researchers should therefore follow their institution's governance procedures and consult their Data Protection Officer where appropriate before commencing processing.

66 *ibid*, arts 44-46.

This reframes local deployment as more than a risk-mitigation tactic. It operationalises privacy by design, embedding safeguards into the processing architecture itself rather than relying on third-party contractual assurances.

4.4. Lawful basis and the research exemption

Compliance is not exhausted by the question of *where* data are processed; it also depends on *why* processing is lawful. As a university researcher undertaking funded work, the applicable lawful basis is as a ‘public task’ under Article 6(1)(e), reflecting the institution’s function as a public body conducting research in the public interest.⁶⁷ The lawfulness of research carried out under Article 6(1)(e) is a narrower, more institutionally anchored basis than the ‘legitimate interests’ ground more commonly relied upon by commercial actors.⁶⁸

5. The Reflexive Local AI (RLAI) Framework

This section develops the RLAI framework that I used. It forms five principles designed to better preserve methodological rigour, participant confidentiality and legal compliance.

67 *ibid*, art 6(1)(e). It is noteworthy that the present project did not involve the processing of ‘special category personal data’ within the meaning of Article 9 UK GDPR. Accordingly, no additional Article 9 condition was required. Nevertheless the governance principles developed in this article may be of particular significance for empirical legal research that does involve special category data, where researchers must also identify an appropriate Article 9(2) condition (for example, Article 9 (2)(j), as supplemented by Schedule 1 of the DPA 2018), in addition to satisfying the obligations discussed in the main text. Where special category data are inadvertently disclosed by a participant despite not being sought, the RLAI framework’s Principle 3 (reflexive contextual anonymisation) extends to redacting content at this stage.

68 *ibid*, art 6(1)(f).

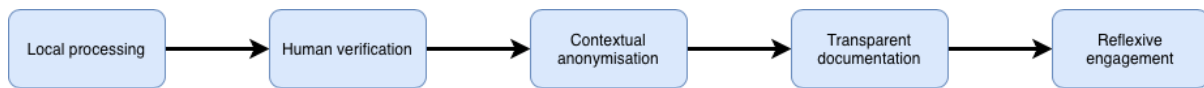


Figure 1 - RLAI model

5.1. Principle 1: Local processing as default

Data sovereignty should be the operational default, not an optional safeguard reserved for unusually sensitive projects. As demonstrated above, offline deployment resolves the Article 28 processor problem, the Article 5(1)(b) purpose-limitation risk and the Chapter V transfer regime simultaneously. It should be noted that this is not because local models are technically superior to commercial alternatives, but because researcher control over the processing environment is a methodological precondition, not merely a legal convenience.

5.2. Principle 2: Human interpretive oversight

No AI-generated transcript should be treated as a final research artefact. The workflow instead proceeds through three stages: (i) an AI-generated pass transcript, (ii) systematic human verification against the original audio, and (iii) only then a finalised transcript suitable for coding. This principle directly addresses the risks of algorithmic deference identified earlier, i.e. the tendency to treat fluent, confidently formatted machine output as self-evidently accurate.

5.3. Principle 3: Reflexive contextual anonymisation

Automated redaction should be understood as a preliminary fix, not a security shield. The researcher, not the model, determines what constitutes an identifier in a given field. As the TCFD vignette in table 1 illustrates, a phrase entirely unremarkable to a generic language model (e.g. a net zero target date, the size of a fund or a specific legacy asset holding) can function as a *de facto* identifier within a small, well-networked elite population. This judgement depends on field-specific knowledge no current model possesses and this cannot be delegated.

5.4. Principle 4: Transparent documentation

Consistent with transparency norms already expected in qualitative methods, researchers should document: (i) the model and version used, (ii) any prompts or configuration parameters applied, (iii) the verification process undertaken, (iv) the corrections made at the human-review stage, and (v) the reasoning behind anonymisation decisions. This creates an auditable record capable of withstanding both peer and regulatory scrutiny.

5.5. Principle 5: Reflexive engagement in practice

Researchers should maintain a parallel reflexive log during processing, capturing what the transcript itself cannot: tonal shifts, hesitation and moments where the model's smoothing of speech may have altered meaning. This principle operationalises resistance to algorithmic deference at the point of practice, distinct from the broader theoretical discussion of automation bias addressed below.

5.6. The framework applied in practice: worked example

The interaction of the five principles is best illustrated by returning to the TCFD vignette introduced earlier. The relevant passage, in which a fictional participant described tying a net zero commitment to a phased divestment timeline while retaining a legacy equity stake, passed through each stage of the RLAI framework as follows.

Principle 1 (local processing) meant that the audio never left the researcher's machine. Whisper generated a first pass transcript entirely offline, so no fund-specific detail was ever exposed to a third-party processor or training corpus. Principle 2 (human interpretive oversight) then required verification of that pass against the original recording; this caught a domain-specific error consistent with the pattern reported earlier in this article, where 'fiduciary' was rendered as a phonetically adjacent but substantively wrong term, requiring correction before the transcript could be treated as usable. Principle 3 (reflexive contextual anonymisation) is particularly important in this example: the model itself had no reason to flag the net zero date or the legacy asset holding in South Wales Oil Co as identifying; this is because neither is a name, institution or postcode. Only field-specific knowledge (that a small number of pension schemes hold this particular legacy position, and that the combination of fund size, sector and timeline is close to unique) allowed the researcher to recognise the passage as a *de facto* identifier and redacted accordingly. Principle 4 (transparent documentation) required logging that redaction decision at the point it was made, together with the model version, the correction to 'fiduciary', and the reasoning for treating the divestment detail as identifying. Of course, this transparency was required so a second

researcher, or indeed a reviewer, could audit the choice rather than take it on trust. Finally, Principle 5 (reflexive engagement) captured something none of the preceding four stages could: a note, made contemporaneously, on the participant's tone when disclosing the holding. There, it was clear that there was a mixture of defensiveness and rehearsed justification that shaped how the passage was later coded, but which the transcript's smoothed, grammatical prose gave no trace of.

No single principle would have been sufficient alone. It is the sequencing and interaction of the five principles, not any one of them, that constitutes the RLAI framework's contribution.

6. Boundaries of the RLAI framework and directions for future research

This section ends the work by looking at the limits of the framework and future research directions that may be taken.

6.1. Limits of the RLAI Framework

Honesty about the framework's boundaries strengthens rather than weakens its contribution.

Three limits are worth stating explicitly.

First, local deployment carries a hardware floor. For instance, this project's own experience of running Whisper's higher parameter models required close to the practical limit of a 16GB unified memory of the system I was using. This is not to say that a researcher cannot use a

standard, university laptop with, say, a minimum of 8GB RAM to run a transcription; but a lesser model could likely limit the accuracy of the outputted final transcript. Thus a framework partly motivated by protecting research budgets from prohibitive transcription costs should not be allowed to quietly reintroduce a different barrier: researchers without access to sufficiently capable hardware face a real constraint of their own. Relatedly, Principle 1's legal benefit is not costless. As section 3.3 and appendix I note, local open-source models such as Whisper measurably underperform on regional accents and offer no out-of-the-box speaker diarisation, so the compliance gain of local processing must be weighed against a real, if manageable, accuracy and functionality cost relative to commercial alternatives.

Second, this project's empirical base is a single-researcher, single-project case study in one substantive field (pensions and climate governance). The RLAI Framework's principles are designed to generalise, but their application has not yet been tested across other socio-legal subfields, multi-research teams or non-English language fieldwork, each of which may surface constraints this project did not encounter.

Finally, the framework presupposes a level of technical competence (installing dependencies, configuring model parameters, troubleshooting command-line errors) that is not evenly distributed across legal (or, for that matter, non-legal) academia. This project benefited from the second author's technical background at the pre-data collection setup stage; researchers without an equivalent resource may face a steeper learning curve than the 'modest consumer-grade' hardware framing above.

6.2. Future research directions

The framework may be extended beyond transcription and into analysis. Researchers have already tested AI in data analysis. For instance, Perkins and Roe undertook a comparison of manual thematic analysis against ChatGPT (GPT-4) assisted coding and found ‘a complementary relationship between GenAI and human researchers’.⁶⁹ This article does not extend its five principles to coding or thematic analysis, but the underlying risk is the same one Principles 2 and 5 are designed to guard against.

Given the RLAI framework promotes AI-assistance, it is important to look at the effect that using AI has on researchers. Thus Christou suggests that delegating interpretive tasks to generative AI risks undermining ‘reflexivity through cognitive offloading’.⁷⁰ Their proposed corrective is a concept called ‘forward reflexivity’, which asks researchers not merely to audit AI output in the moment, but to anticipate how AI-augmented research texts may themselves circulate and return as training data for future models.⁷¹

69 Mike Perkins and Jasper Roe, ‘The use of Generative AI in qualitative analysis: Inductive thematic analysis with ChatGPT’ (2024) 7(1) *Journal of Applied Learning & Teaching* 390.

70 Prokopis A Christou, ‘Reconfiguring Reflexivity in the Era of AI: From “Turning Back” to “Looking Forward” Through Constructivist and Posthumanist Lenses’ (2026) *Qualitative Inquiry* 1-10.

71 *ibid.*

And finally, the project's local processing approach should connect the broader literature on 'digital sovereignty', which Roberts sees as a normative concept.⁷²

7. Conclusion

The intersection of generative artificial intelligence and empirical legal research presents a profound methodological paradox. While automated transcription offers powerful solutions to qualitative bottlenecks, transcription remains an inherently constructivist act that actively shapes all subsequent legal analysis. Outsourcing this interpretive layer to opaque cloud infrastructure risks violating strict data protection mandates under the UK GDPR and DPA 2018, compromising elite participant confidentiality and introducing automation bias into the workflow.

The RLAI Framework resolves these challenges by anchoring data sovereignty and human-led critique at the centre of empirical research. By enforcing local execution, multi-stage verification, contextual pseudonymisation, and transparent documentation, the framework proves scholars need not choose between methodological rigour and technical efficiency. Ultimately, preserving the integrity of legal scholarship in the AI era requires technology to serve as a secure instrument that reinforces, rather than replaces, critical human judgment. In this sense, the RLAI framework is offered not merely as a transcription protocol, but as a template for the governance of local AI in empirical legal research more broadly.

72 Huw Roberts, 'Digital sovereignty and artificial intelligence: a normative approach' (2024) 26 Ethics and Information Technology 70. See also House of Commons Library, Digital sovereignty (CBP 10547, 2026) <<https://commonslibrary.parliament.uk/research-briefings/cbp-10547/>> accessed 1 August 2026.

Appendix I: Technical guidance on using localised AI models

This appendix sets out the technical workflow underpinning the study.

Hardware

The transcription processing was performed locally on an Apple Mac mini (M4, 2024; 10-core CPU, 10-core GPU, 16GB unified memory). This represented a modest, consumer-grade baseline. This hardware configuration comfortably accommodates Whisper's full range of model sizes; as mentioned within the software, memory pressure remained within practical limits of the unified memory utilised.

The 16GB unified memory ceiling constrained language model selection. That said, this baseline configuration successfully demonstrates that modest consumer-grade hardware can achieve optimal results for local transcription workflows. While larger parameter models could theoretically offer enhanced context awareness, the M4 Mac mini set up proves sufficient for robust, secure execution under the RLAI framework.

It is important to note that when running Whisper, foreground applications were closed to maximise available memory during transcription; background system services and processes managed by macOS itself, which cannot be user-terminated, were permitted to continue running as normal

Software and model selection

Whisper

OpenAI's Whisper ASR model was used because it is open source and provides the ability for transcribing audio in various formats, e.g. '.wav' and '.mp3'.⁷³ It is also compatible with the recording tools used, as mentioned below.⁷⁴ The model has been compiled from a large dataset of audio recordings, which allows specifically for speech translation and language identification, making this an appropriate tool for transcribing interview recordings.⁷⁵

Whisper segments the transcription logs which provides further details and information on the timestamp of the chats; it can even provide granular results of timestamps per word if needed (captured as part of a 'VTT file' – a text-based subtitle/caption file). However, the use case for this research was to ensure timestamps per sentences and this is recommended for qualitative interview transcripts.

OpenAI Whisper provides six model sizes to work with.⁷⁶ The models can be English-only or multilingual; the model sizes range from 'tiny' to 'turbo'.⁷⁷ Whisper documentation recommends using English-only models which offers speed and accuracy (measured by transcribing English speech on a A100 GPU), but is limited only up to medium.en (769m parameters) up to ~2x the speed.⁷⁸ The multilingual models allow up to turbo (809m parameters) at a much faster processing speed of ~8x.⁷⁹ This latter model was selected for use with the Mac mini as a result of its speed of operations. Whilst a multilingual model was used,

73 OpenAI, 'Whisper' (GitHub, 2022) <<https://github.com/openai/whisper>> accessed 1 August 2026.

74 *ibid.*

75 *ibid.*

76 *ibid.*

77 *ibid.*

78 *ibid.*

79 *ibid.*

English can be configured as a way for transcribing as part of the tool.⁸⁰ More information on the performance and how these models have performed against different languages can be found on their GitHub repository [OpenAI/whisper](#).⁸¹ There is no known telemetry which is collected from using or running Whisper.

A limitation of Whisper is the inability of knowing which person said what sentences, which is called diarisation. Nonetheless the ability to diarise the transcript can be accomplished by using `pyannote` (a Python library) and other open-source audio speaker diarisation models. However, this workflow focussed primarily on transcription and timestamping, since the context of the transcriptions were more important.

This guidance does not instruct on how to download and install Whisper (instructions can be found in the official GitHub webpage).⁸² Once installed, here is an example command in a ‘terminal’ of how to run the whisper toolset with an example .mp3 audio file. The below example takes a .mp3 audio file and setting to an English-only model, ensuring the language for transcribing is set to English. To add timestamps to the transcription the `--word_timestamps` parameter set to `True` adds this to the output. Setting the output format will only provide the formats set, otherwise if not, then all associating files including the .vtt (for text-based subtitling) will be available. The below command, once complete, will output

⁸⁰ *ibid.*

⁸¹ *ibid.*

⁸² *ibid.*

the transcription and create two new files. One which output the plain English transcribed text (from the `--output_format` flag): `interview1.txt`, and the other including the time stamped sentences from the redirected output, `interview1_timestamp.txt`.

```
whisper interview1.mp3 --model turbo --language en --task transcribe --  
output_format txt --word_timestamps True > interview1_timestamp.txt
```

Python

Python allows libraries like Whisper to be used with greater flexibility than the command-line interface alone, providing more advanced control over processing and output. The specific scripts are not included here as they are Whisper's core functionality wrapped in a bespoke interface rather than a generalisable tool, but the approach can readily be reconstructed using the Whisper documentation cited below. We created our own set of Python scripts which handled the running of Whisper, using segmentation of sentences to build and output `.txt` files (with timestamps) into organised folders with clear transcribed files.

Recording and file handling

Interviews were recorded using a handheld Sony IC Recorder (model: ICD-PX370). The device has no wireless or cellular connectivity, meaning the audio file could not be transmitted or synchronised to any cloud service at the point of capture, consistent with Principle 1 of the RLAI framework (local processing as default).

The recorder's built-in USB-A connector allowed audio files to be transferred directly to the researcher's local machine. Transcription was then performed entirely locally, using Whisper

(as described above). Audio files and completed transcripts were stored on the researcher's Swansea University's password protected OneDrive account, governed by the university's institutional data processing agreement with Microsoft. As already established, this is distinct from the consumer facing AI platforms discussed in the article. Microsoft's enterprise terms for institutional OneDrive accounts do not include the training-data clauses identified as the central Article 5(1)(b) concern in this article, and the university, not Microsoft, remains the data controller. As a backup, the 'British Academy/Leverhulme Project' OneDrive folder was regularly backed-up onto a USB-C password-protected flash drive and stored in a locked desk drawer.

The pseudonymised Microsoft word documents containing the transcripts were password protected. When a participant's transcript was sent via email, the transcript would be attached and sent in one email and the password in another, separate email. However, in most cases, the password was sent separately to participants using LinkedIn's private messenger service.

Verification process

Consistent with Principle 2 of the RLAI framework (human interpretive oversight), no output generated by Whisper was treated as a finished transcript. Each recording produced an AI-generated first pass, which was then checked in full against the original audio. During transcription and pseudonymisation, the computer's WiFi was turned off, to ensure that the work remained entirely local.

The verification process required listening to the entirety of each recording alongside the corresponding AI-generated text, correcting errors as they arose. As discussed in section 3.2 above, the most consistent category of error concerned domain-specific socio-legal terminology, e.g. the term ‘fiduciary’. As the transcripts were well done, a rough estimate of how long the verification for a one-hour interview generally took would be around three hours. This time would have been shorter, but I liked to be thorough, and the transcript would be checked against the audio on ‘normal’ speed first and then later on the ‘fast’ speed setting.

By correcting misrecognised words, this stage also served the contextual pseudonymisation described in Principle 3: identifying and redacting details that, while not explicit identifiers such as names or institutions, could reasonably function as identifiers within a small, well-networked elite population (see the worked example above). Redaction decisions of this kind were recorded in a documentation log, consistent with Principle 4, together with the model version used and the corrections made at this stage. Only once this verification and pseudonymisation was complete was a transcript treated as finalised and suitable for coding and analysis.