Client Data, AI, and the Boundaries of Consent: A Practical Note for Industry Leaders

Following the issuing of the Insights Association's statement about the re-use of research data to train AI models, MRS and Esomar PSC have agreed the following joint practical guidance note.

As AI tools become embedded in everyday research and data workflows, a question that was once considered settled is back on the table: who can do what with the data, materials, and learning generated by a project? Historically, our sector has operated on the clear assumption that client data (e.g. transcripts, notes, draft outputs, interim analyses) belongs exclusively to the client, and that anything an agency or supplier does with it is bound by that principle. That assumption is now being tested, often quietly, in vendor terms of service, platform updates, and AI feature rollouts. If this is becoming a topic of conversation, it needs to become a topic of explicit agreement. The specific question this note addresses is the use of client and participant data to train AI models and processes, rather than the use of AI in research more generally.

This note is intended for Insight Managers, agency leaders, and senior players across the supply chain. It is not legal advice. It is a practical prompt to think about what you are committing to, what you are entitled to commit to, and what your contracts, undertakings, and platform agreements actually say. Underpinning all of it is participants' consent: they agree to a specific purpose, and that consent is the upstream constraint on everything done with the data afterwards.

It is worth being clear about the scope. The actors in question are the intermediaries (platforms, vendors, agencies, and sub-processors) that handle data on behalf of clients and participants, and the issue is what they may or may not do with it. This note does not seek to constrain what end-clients do with their own data, provided that use remains consistent with the permissions given by research participants. Its focus is the intermediaries, such as panel companies, recruiters, and software platforms, that handle data on behalf of clients and participants These questions apply to qualitative work as much as to quantitative: focus group recordings, depth-interview transcripts, ethnographic notes, and community board content can be just as exposed as survey data, and often more revealing of client thinking. They also apply to platforms the industry has used for years, not only to the latest generation of AI services. If focus group transcripts are uploaded to a translation platform, an automated transcription service, a cloud storage tool, a survey platform, an analytics dashboard, or a project management system, the same questions arise: what do the terms permit, are those materials being retained, and are they being used to train the supplier's models? The introduction of AI features into established services means that platforms cleared years ago may now be doing something quite different with the same data. Old contracts deserve a fresh read.

1. Restate the default, in writing

If your working assumption is that client data and the insights derived from it remain the property of the client, say so plainly in your contracts. The same principle applies in reverse: if a client engages you, they should be able to point to a clause that confirms it. Silence is no longer safe. Vendors, platforms, and AI suppliers are increasingly relying on broad language (“internal analyses”, “benchmarking”, “aggregated and de-identified use”) that can quietly cover model training, derivative outputs, and reuse across unrelated clients.

2. Know what you can, and cannot, consent to

This is the issue most likely to catch suppliers out. An agency may be asked by a platform to agree to a particular use of project information. Before clicking accept, ask:

- Do the research participants understand and permit this use? Their consent was given for a specific purpose.
- Does the end client permit it? Their contract may prohibit downstream reuse, even in anonymised or synthetic form.
- Do upstream partners, panel providers, recruiters, or sub-processors permit it? Their terms flow through to you.

You cannot grant rights you do not hold. A signature on a vendor T&C does not retroactively create permissions from people who were never asked.

3. Separate the two things that anonymisation does, and does not, do

Two distinct obligations lie beneath this type of activity, and they are often conflated when AI is involved. They need to be treated separately, because they are owned by different parties and protected by different means.

Personal data, owned by participants. Removing names, contact details, and other identifiers protects the people who took part in the research. This is a privacy obligation, governed by data protection law and by the consent participants gave at the point of collection. Anonymisation, properly done, addresses this layer, although if confidentiality assurances are given to participants, such as restricting the use and dissemination of any research data, obligations to restrict the use of anonymised participant data will remain.

Client's intellectual property, owned by the client. The questions a client chose to ask, the hypotheses they were testing, and the patterns and findings those questions revealed are commercial assets the client paid to obtain. Anonymising participants does nothing to protect this layer. A de-identified dataset, an aggregated summary, or a synthetic output can still carry the client’s strategic thinking and the answers they paid to discover. If those outputs feed a model that is then available to unrelated third parties, the client may reasonably feel they have funded a competitor’s advantage.

The practical implication: anonymising participants does not anonymise the research. Anonymisation may discharge the participant-facing duty but leaves the client-facing duty fully in place.

It is also worth noting the need to apply the terms anonymisation, de-identification, and pseudonymisation consistently, appropriately and not interchangeably, since the distinctions matter

both legally and practically. Demonstrable and robust anonymisation is that which reduces the likelihood of identifiability to a remote level, and this needs to be applied at every stage of the data lifecycle.

4. Look downstream as well as upstream

Most agencies and suppliers pay close attention to what they promise their clients. It is also essential to scrutinise what their tools, platforms, and AI vendors are entitled to do with the same data. The two need to match. Where vendor terms reserve broad rights for "service improvement" or "model training", the agency may be in breach of its client undertakings even when the vendor is acting entirely within its own. Read the vendor terms with your client commitments in the other hand.

5. Make disclosure the norm

Clients increasingly need to know which AI tools have touched their project, at what stage, and for what purpose. Client contracts may also require advance notice/approval before using subcontractors or additional processors. A short, standing disclosure (which platforms are used, for what, and on what terms) is becoming a baseline professional expectation, and is a requirement within some of the sector's ethical Codes. It is also a useful internal discipline: if a use is hard to disclose, that is usually a signal.

6. Build an internal framework that matches actual work

"Use AI responsibly" is not a policy. A practical AI framework should cover: what data may be entered into which tools, who approves harder cases, when clients should be told, and how decisions are recorded. Training should use real examples from research practice (transcripts, open-ends, synthetic profiles, AI note-takers, deck drafting), not abstract principles. There should be minimum controls put in place which can increase efficiency when using such frameworks e.g. approved tools lists, data classification (e.g. what is permitted with approval, what is never permitted, etc), approval/escalation routes, data retention and deletion requirements.

A shared call to the industry

MRS, Esomar and the Insights Association Codes and related guidance are clear that protecting participant and client trust is not optional as AI reshapes research practice. Practitioners and organisations, particularly the intermediaries who handle data on behalf of others, need to establish an approach which satisfies their ethical and legal obligations, including: stating ownership and permitted use clearly in every contract; ensuring that what is agreed with vendors is consistent with what has been promised with clients and participants; and treating anonymised, aggregated, and synthetic outputs as still bounded by the original commitments under which the data was collected.

The historical default that client data belongs to the client remains a sound starting point. What has changed is that it can no longer be left unsaid. This is not necessarily good or bad, but it needs to be acknowledged and dealt with.

To find out more

MRS, Esomar and the Insights Association have guidance on using AI and this should be referred to find out more about your ethical and legal obligations:

MRS Guidance on Using AI and Related Technologies

ICC/Esomar Code on Market Research, Opinion and Social Research and Data Analytics

Also see the Insights Association assessment of the legal issues concerning client research data re-use for AI and the Insight Association's Code of Standards & Ethics for Market Research and Data Analytics.