AI data-use contract terms checklist

For contracts with AI providers (vendor APIs, cloud AI platforms, AI features in research software) and with partners who may use AI on your research data. 17 items. Published terms checked 30 October 2026; negotiated agreements can differ and terms change. Issues to check with counsel, not legal advice.

1 Training use

- ☐ The agreement itself (not only a help page) says the provider will not use your content to train or improve its models. [1,2,3]
- ☐ Exceptions are known and set: feedback buttons, opt-in data sharing, 'service improvement' wording. Off for research work unless decided otherwise. [4,1]
- ☐ If training is ever allowed: opt-in only, confidential and proprietary data excluded, training stops at termination. [5,6]

2 Retention and deletion

- ☐ A retention period for each category: prompts, outputs, logs, files, fine-tuning data, metadata. No 'reasonable period'. [6,7]
- ☐ A deletion deadline after termination, and its carve-outs (law, disputes, abuse or harmful use). Published: OpenAI 30 days; Anthropic 30 days; Microsoft 90 + 90 days; Google 30-day recovery + up to 180 days. [1,8,9,10]
- ☐ Any clause allowing the provider to anonymise and keep your data for research or statistics is found and, for research data, usually excluded. [11]
- ☐ Features that keep data until you delete it (stored state, file stores, tuning files) are listed and owned. [7,12]

3 Derived or fine-tuned models

- ☐ A model tuned on your data is for your use only and can be deleted with its training files. [7,13,3]
- ☐ Definitions of usage, aggregated and 'derivative' data do not let learnings from your data reach the vendor's general models. [5,14]
- ☐ Restrictions on training your own models on outputs are understood before you plan to do it. [1,3]

4 Licence and sub-processors

- ☐ You keep inputs and own outputs; any licence back is limited to delivering the service, time-limited and explicit. No perpetual licence to improve the tool. [1,2,6,15]
- ☐ Sub-processor list, notice period and right to object are known, including whether they cover non-personal research data (OpenAI and Anthropic DPAs cover personal data). Microsoft 6 months' notice (30 days for AI sub-processors); Google 30 days, objection by terminating. [16,8,9,10]
- ☐ Equivalent data-protection obligations flow down to each of the provider's sub-processors. [6,16,8]

5 Flow-down to collaborators and CROs

- ☐ Each outbound CDA, CTA, CRO or collaboration agreement says whether the partner may put your confidential information into AI tools, and on what terms. [17,18]
- ☐ Outputs made from your information count as your confidential information, and the restriction binds the partner's representatives. [17]
- ☐ You may ask the CRO or vendor which AI tools it uses, and assess them in advance; the sponsor stays responsible for delegated trial activities. [19,20]
- ☐ Inbound agreements (CDAs, MTAs, collaborations) are checked for limits on processing received information with AI; outside agreed conditions the owner's consent may be needed. UK NHS model agreements are used unaltered: plan around them. [17,18]

Also: keep a dated copy of the terms you rely on and re-check on notice of changes (changes required by law can apply sooner) [1,2]. Your contracts and controls can be evidence of the 'reasonable measures' US trade-secret law asks of an owner [21]. Patent, export-control, personal-data and US bulk-data (28 CFR 202) questions go to counsel.

Sources 1. OpenAI Services Agreement (eff. 1 Jan 2026) 2. Anthropic Commercial Terms of Service 3. Google Cloud Service Specific Terms 4. Anthropic Privacy Center: Is my data used for model training? 5. Morgan Lewis: Key Considerations When Allowing a Vendor to Train Its AI Models on Customer Data (Jun 2026) 6. Tucker Ellis LLP: The AI Vendor Contract Is Becoming More Important Than the Privacy Policy (Aug 2026) 7. OpenAI Enterprise privacy 8. Anthropic Data Processing Addendum 9. Microsoft Products and Services DPA (May 2026) 10. Google Cloud Data Processing Addendum 11. Anthropic Privacy Center: How long do you store my organization's data? 12. OpenAI API docs: Data controls in the OpenAI platform 13. Microsoft Learn: Data, privacy, and security for Foundry Models sold by Azure 14. Loeb & Loeb: Navigating the Challenges of AI Contract Negotiations (May 2026) 15. Bristows: Integrating generative AI into the life science ecosystem 16. OpenAI Data Processing Addendum 17. Trethowans: NDAs in the new world of AI 18. Health Research Authority: model agreements and AI use of confidential information (7 Sep 2026) 19. Freshfields: Hot Topics in Global Clinical Trial Agreements (ACC, Mar 2024) 20. ICH E6(R3) Guideline for Good Clinical Practice (Jan 2025) 21. 18 U.S.C. 1839, definition of trade secret (Cornell LII)

Atlas Bio Systems (atlasbiosys.com). Free to copy and use inside your organization. Full explanation: https://atlasbiosys.com/answers/keep-proprietary-research-data-out-of-external-ai-models.html . Latest version: https://atlasbiosys.com/downloads/ai-data-use-contract-terms-checklist.pdf