instats

P1 — Model Institutional AI-in-Research Policy

An adopt-and-adapt policy text for the use of artificial intelligence in research and research training, drafted so that a general counsel and a research office can take it into governance without starting from first principles.

Browse the full online instrument catalogue to read or download every resource in HTML, Word, PDF, and spreadsheet formats.

About this instrument

Purpose	To give an institution a complete, adoptable policy covering AI use in research and research training, reaching beyond basic plagiarism into valid research practice, competency, and examination.
Who completes it	The research office and general counsel jointly, together with the graduate school and the research-integrity function. Approved by [COMMITTEE].
Time required	Two to three drafting sessions to work through the local decisions listed at the end of this document, plus one governance cycle for formal approval. An institution that already has an AI policy can adapt Clauses 5–9 and 13 while retaining its own Clauses 1–3.
Related report sections	Part 1.3 (the four policy classes), Part 2 (all five dimensions), Part 4 and Appendix A (the maturity grid), Appendix C (the eleven AI tool classes), and Appendix G (the labor-and-judgment task taxonomy).
Cells touched	Adopting this instrument supplies evidence toward the **policy** cells for all five dimensions (D1-policy through D5-policy). Clause 13 identifies the specific operational owners whose supporting artifacts A1 evaluates for the **people** cells (D1-people through D5-people). The policy also specifies record and control requirements relevant to **D2-systems**, **D3-systems**, **D4-systems**, and **D5-systems**, along with operating requirements relevant to the **process** cells in Clauses 9–12 and 15. This instrument does not directly supply evidence toward **D1-systems**, because the systems that enforce the line between labor and judgment operate inside local advising, supervision, and examination workflows rather than inside a policy text. No cell or dimension advances automatically. A fresh A1 assessment determines cell levels, and a rescore across all four cells determines each dimension's level.
Ceiling	This policy supplies evidence toward the *established* rung at most for the named cells. It does not itself establish a cell or dimension. Reaching the *leading* level requires published scoring, documented outcome evidence, and external benchmarking (*Responsible AI in Academic Research: A Competency Framework for Research Training* §4.1), which are supported by instruments A1, A4, and A5 in this pack.

Table 1. Overview of the policy instrument, outlining its purpose and key administrative characteristics.

How to use this template

The companion report classifies institutional AI policies across thirty-eight top-tier doctoral universities in fifteen countries and jurisdictions into four distinct classes based on their scope. Class A simply appends an AI clause to an existing student-conduct policy (four institutions, approximately 11 percent). Class B creates a dedicated AI policy that stops at plagiarism and co-authorship rules (eleven institutions, approximately 29 percent). Class C extends policy coverage into research integrity, reproducibility, and formal disclosure (seventeen institutions, approximately 45 percent), which is currently the most common approach across universities. Class D adds codified competency expectations for doctoral students and supervisors, explicit curriculum requirements, and clear rules for examiners (six institutions, approximately 16 percent). In United States usage, examiners are the members of the dissertation committee. You can find the full distribution and analysis in the report §1.3 and §3.2.

I have written this template at Class D. Key elements of this template draw directly from observable practices in the report's university sample (§1.3, §3.2), and the drafting notes for Clauses 5, 10, and 11 provide the underlying details:

- AI disclosure integrated into the German sworn submission declaration (the *eidesstattliche Versicherung*) at Heidelberg's Graduate Academy, which is the German cohort's single Class D policy.
- Examiner-side conduct rules from the two UK Class D documents, developed by University College London's Doctoral School and King's College London's Centre for Education Studies.
- The requirement that a doctoral student must be able to describe and defend any use of generative AI, as well as the overall contents of the dissertation, during the final oral examination or defense, adapted from the University of Toronto School of Graduate Studies, which the report classifies at **Class C**.
- A codified research-training learning path at KU Leuven and a formal literacy requirement tied to institutional course offerings at the University of Helsinki, the two continental European Class D institutions outside Germany.
- Tsinghua University's December 2025 institution-wide framework, which is the only policy in the sample that covers teaching, research, and graduate theses in a single integrated document.

If your institution prefers to adopt a Class C rather than a Class D policy, you can adopt this entire template while simply omitting Clauses 10.2–10.4 (the codified competency and training requirements) and Clauses 11.4–11.7 (the examiner-side rules). Those two specific areas are what define Class D in the report's taxonomy. Clause 11.2 can remain in place either way, since the report documents that defensibility rule at a Class C institution. Choosing Class C policy scope can be a valid institutional decision, but the Class A–D scope taxonomy is distinct from the maturity grid. That choice should be a deliberate, recorded governance decision rather than an accidental drafting omission.

Reading the annotations. Every clause in this template includes explanatory annotations. The **drafting note** explains why the clause is structured the way it is and highlights what you should adjust for your local context. The **framework key** identifies the maturity grid cells supported by the clause, the report section it draws from, and the concrete artifact that would demonstrate to a skeptical auditor that the policy is actively operating in daily practice rather than just sitting on a website. Where relevant, **jurisdictional variation** notes explain where the wording needs to be tailored to specific regional or national legal requirements.

Variable fields appear throughout the template in square brackets, such as [INSTITUTION NAME], [COMMITTEE], [ROLE], [DATE], and [N]. Please delete all explanatory annotations before publishing your final policy document, because they are intended solely for the drafting and governance teams.

Policy on the Use of Artificial Intelligence in Research and Research Training

[INSTITUTION NAME]

Policy owner	[ROLE (recommended: Deputy Vice-Chancellor (Research), Vice-President or Vice-Provost for Research, or equivalent)]
Approving authority	[COMMITTEE]
Approved	[DATE]
Effective	[DATE]
Next scheduled review	[DATE]
Version	1.0
Supersedes	[DOCUMENT], [DATE]

Table 2. Administrative metadata and designated governance roles for the institutional research AI policy.

Clause 1 — Purpose

1.1 This policy establishes [INSTITUTION NAME]'s official position on the use of artificial intelligence in research and research training. It defines which research judgments must always remain human, explains what responsible AI use means at the level of specific research tasks, specifies what must be disclosed, sets rules for which tools can be used with different classes of data, clarifies the responsibilities of advisors, supervisors, and examiners, and establishes institutional accountability.

1.2 This policy is not an academic-integrity policy and does not replace one. Academic-integrity policies govern assessed student coursework. This policy governs the conduct, reporting, and dissemination of research, including research conducted for graduate research degrees.

1.3 This policy exists because the phrase "responsible AI use" appears across funder guidelines, journal policies, and university statements without carrying much operational meaning on its own. For the purposes of this policy, responsible use means AI use that is simultaneously **ethical, valid, reproducible, and transparent** (the report §1.2). All four properties are required together.

> ***Drafting note.*** *Clause 1.2 is essential and represents the structural piece most often missing in the policies surveyed in the report. Fifteen of the thirty-eight universities in its sample (roughly two out of five) stop at a plagiarism-policy baseline, and four of those place AI entirely within student academic-integrity rules (§1.3, §3.2). Under that setup, an invented citation in a faculty member's grant application has no clear institutional governance home, whereas the exact same fabrication in a student's coursework triggers a formal integrity hearing. If your institution already has an AI clause within its student academic-integrity rules, cross-reference it here and state clearly that this policy governs all research outputs.*
>
> *Framework key: D1-policy, D2-policy (the report §1.2, §1.3, §2.1). Evidence of operation: a published policy document with a recorded version date and the approving committee named in the official minutes.*

Clause 2 — Scope and application

2.1 This policy applies to everyone engaged in research activities under the auspices of [INSTITUTION NAME]. This includes academic and research staff, doctoral and other graduate research students, research advisors and supervisors (and members of supervisory panels), internal and external examiners, postdoctoral and early-career researchers, visiting and honorary scholars, professional staff who support research activities, and any external contractors or consultants engaged in research.

2.2 This policy applies to all research activity conducted under the name of [INSTITUTION NAME] or using its facilities, equipment, or resources, regardless of the funding source, academic discipline, or whether the work is intended for formal publication.

2.3 This policy applies to AI use across every stage of the research lifecycle, including research design, ethics review, data collection, data analysis, interpretation of results, drafting, submission, examination, publication, peer review, and public dissemination.

2.4 Where this policy conflicts with a binding external requirement (such as a research funder condition, a publisher policy, a contract term, a data-sharing agreement, or a statutory legal obligation), the stricter requirement applies and the researcher must comply with it. The [ROLE: Research Office] maintains the official register of external requirements in Schedule D.

2.5 The existence of a scientific-research exemption in an applicable AI or data-protection statute does not waive or override this policy. Statutory research exemptions protect different legal interests across different jurisdictions, and none of them addresses or guarantees research validity.

> ***Drafting note.*** *Clause 2.5 closes a common legal loophole that researchers and counsel often raise. Four major legal instruments include scientific-research exemptions, but each covers something different: the EU General Data Protection Regulation Article 89 (a conditional derogation framework), the EU AI Act Article 2(6) (a categorical exclusion that Article 2(8) qualifies by keeping real-world testing in scope), the Council of Europe Framework Convention Article 3(3) (the narrowest carve-out), and the UK Data Protection Act 2018 Schedule 2 Part 6. A project exempt under one statute is not necessarily exempt under another, and none of these laws speaks directly to research validity. When finalizing Clause 2.4, decide whether individual schools and faculties may adopt stricter local rules (I recommend yes) or looser ones (I strongly recommend no).*
>
> *Framework key: D1-policy, D5-policy (the report §2.5). Evidence of operation: the policy's own scope clause, along with the Schedule D register showing a recent review date.*

Clause 3 — Definitions

3.1 **Artificial intelligence system** means a machine-based system that, for a given set of explicit or implicit inputs, generates outputs such as text, predictions, recommendations, classifications, or other content. **Generative AI** refers to any AI system whose primary function is generating new content.

3.2 **Agentic research tool** means an AI system that executes multi-step research tasks to produce outputs that are incorporated directly into the research record. In this policy, agentic tools sit in the same functional family as statistical software packages, automated transcription engines, or data-cleaning scripts. This policy treats AI outputs as tool results rather than as independent authorship (the report, Appendix C).

3.3 **The four AI-use modes.** All AI use in research falls into one or more of four operational modes, and the rules in Clause 6 are organized around them:

Mode	Definition
Search	Using an AI system to scope literature, discover data sources, or locate prior research.
Co-author	Using an AI system to draft, expand, summarize, paraphrase, translate, or edit text or code.
Validator	Using an AI system to check claims, identify errors, critique arguments, or run an adversarial review over a researcher's own work.
Tutor	Using an AI system to explain difficult concepts, work through sample problems, or support individual learning.

Table 3. Definitions and operational descriptions for the four primary modes of research AI use.

3.4 **The eleven AI tool classes.** The specific tool class a researcher selects determines the primary failure mode they must actively guard against. This policy adopts the functional taxonomy set out in Appendix C of the report:

Class	Function in research	Primary failure mode to guard against
1: LLM search	Scoping and discovering literature, with or without live web retrieval	Fabricated citations and misattributed sources
2: LLM co-author	Drafting, paraphrasing, summarizing, and editing prose	Generating substantive claims the researcher cannot defend, or introducing invented terminology into the literature
3: LLM validator	Checking claims and structuring critiques	Missing errors generated by the same underlying model parameters, or displaying unearned confidence in incorrect assertions
4: LLM tutor	Explaining concepts and supporting learning	Domain inaccuracies when working without a verified source corpus, or the student outsourcing the core intellectual tasks the training program is designed to build
5: LLM coder	Generating, debugging, and refactoring research code	Security vulnerabilities, hallucinated libraries and APIs, or code that only passes a single narrow test case
6: Retrieval-augmented research assistants	Answering questions against a specific corpus selected by the researcher	Hallucinations persisting at lower but non-zero rates, or misattributing claims to real sources in the corpus
7: Automated literature-synthesis tools	Prioritizing, screening, and extracting data for systematic reviews	Inconsistent recall across different topics, or inappropriately substituting automated screening for difficult borderline inclusion decisions
8: Specialized research analysis assistants	Domain-specific prediction and analysis (such as molecular structure, forecasting, or material properties)	High-confidence predictions that are inaccurate at the exact resolution required by the research project

Class	Function in research	Primary failure mode to guard against
9: Image and figure generation	Creating illustrative graphics, schematics, and presentation visuals	Any generation or modification of empirical data figures
10: Voice and transcription	Transcribing interviews, focus groups, and field observations	Invented or altered transcript segments, and systematic error or bias against accented or disfluent speech
11: AI-detection tools	Attempting to flag text as AI-generated	High false-positive rates, and documented bias against writers whose first language is not English

Table 4. Research functions and primary failure modes associated with eleven artificial intelligence tool classes.

3.5 **Labor task** means a research task where AI assistance is permitted because the underlying work can be independently verified and recovered by the researcher. **Judgment task** means a core intellectual task where substituting AI for human reasoning silently compromises research validity. Schedule A provides the full institutional list distinguishing labor tasks from judgment tasks.

3.6 **Doctoral student** means any person enrolled in a research doctorate at [INSTITUTION NAME], at any stage from initial enrollment through final degree conferral. Equivalent local titles include *PhD student* or *graduate student* (United States and Canada), *doctoral candidate* or *higher degree by research candidate* (Australia), *postgraduate researcher* or *research student* (United Kingdom), and *Doktorand* (Germany). Where an institution uses *candidate* strictly for students who have passed a formal confirmation milestone, this policy applies more broadly by binding every student from their first day of enrollment.

3.7 **Substantive use** means any AI use that directly shapes what the research record asserts, including content, structural logic, empirical evidence, data analysis, theoretical interpretation, or the selection of primary sources. In contrast, **incidental use** refers to basic mechanical aids such as spellchecking, single-word autocomplete, standard grammar corrections on a researcher's own draft, or automated reference formatting.

3.8 **Approved tool** means an AI tool that is officially listed on the institutional register in Schedule B through the evaluation process set out in Clause 9. **Tool tier** refers to the security and deployment tiers defined in Clause 8.4.

3.9 **Research output** means any item that enters the academic record, including dissertations and theses, journal manuscripts, preprints, conference papers, shared datasets, software repositories, grant proposals, ethics applications, funder reports, and public research communications.

> ***Drafting note.*** *Establishing clear use modes and tool classes is what makes this policy practical on the ground. Simply telling faculty and students to "use AI responsibly" gives an advisor very little actionable advice for a weekly supervision meeting. In contrast, saying "when using validator mode, run your draft through a different model family because self-critique tends to miss its own generation errors" gives them a concrete, practical instruction. Clause 3.7 is often where local debates focus. If your institution has a large multilingual graduate cohort, you should define the boundary for incidental use explicitly rather than leaving it to ad hoc dispute.*
>
> *Framework key: D2-policy, D3-policy (the report §2.2, §2.3, Appendix C). Evidence of operation: the definitions appear verbatim in graduate research training materials and in the institutional disclosure template.*

Clause 4 — Principles

4.1 **Accountability is human and cannot be transferred.** Researchers remain fully accountable for every element of their research outputs, including any content generated with AI assistance. An AI system cannot bear professional responsibility and cannot hold any role that requires it.

4.2 **Verification precedes incorporation.** No AI-generated factual claim, reference, quotation, statistical finding, or transcript excerpt may enter a research output until a named human researcher has verified it against primary source material.

4.3 **Judgment cannot be delegated.** The core research judgments listed in Schedule A must remain human. AI tools may help organize options or scaffold a decision, but they cannot make the decision.

4.4 **Transparency is proportionate to reliance.** The more an output relies on AI tools, the more detailed the disclosure must be. Disclosing AI use is a standard description of research methodology, not an apology for using modern tools.

4.5 **Data confidentiality governs tool choice.** The sensitivity and classification of the data determine which tools a researcher is permitted to use. This is a binding institutional rule, not an individual researcher preference.

4.6 **Tooling is procured on evidence, not vendor claims.** The institution evaluates candidate AI tools against observable operational criteria before approving them, and does not accept vendor marketing claims as proof of performance.

4.7 **Competency is an institutional responsibility.** Doctoral students cannot develop research competencies that their advisors do not understand or model. The institution must provide appropriate training rather than treating AI literacy as an assumed personal trait.

4.8 **The institution evaluates its own progress.** [INSTITUTION NAME] assesses itself against the benchmarking framework on a regular schedule, assigns clear owners for each operational axis, and reports findings annually to [COMMITTEE].

> ***Drafting note.*** *These eight principles correspond directly to operative clauses later in the text. I recommend resisting the urge to add vague aspirational slogans (such as "promoting innovation" or "driving excellence") that have no enforcement mechanism in the policy. When a policy includes unenforced principles, it weakens the credibility of the entire document. Principle 4.6 often faces pushback from research groups that have already adopted their own preferred tools, but Clause 9.6 provides an exception pathway to handle legitimate specialized needs smoothly.*
>
> *Framework key: D1-policy, D2-policy, D3-policy, D4-policy, D5-policy (the report §2.1–§2.5). Evidence of operation: each principle links directly to at least one operative clause in this document.*

Clause 5 — The human-in-the-loop obligation

5.1 [INSTITUTION NAME] maintains and publishes an institutional schedule that distinguishes labor tasks from judgment tasks across the research lifecycle (Schedule A). Individual schools and departments may add discipline-specific tasks to this list, but they cannot remove items.

5.2 **Judgment tasks must never be delegated to an AI system.** A researcher may use an AI system to support a judgment task (such as brainstorming alternative interpretations, identifying potential counterarguments, or structuring a comparative analysis), provided that a named human

makes the final decision, can explain the intellectual justification independently of the AI output, and takes personal ownership of the choice.

5.3 A labor task may be assisted by AI tools, subject to the verification rules in Clause 5.4 and the disclosure requirements in Clause 7. Classifying an activity as a labor task does not permit a researcher to skip either verification or disclosure.

5.4 **The verification obligation.** Before any AI-assisted element is incorporated into a research output, a named human researcher must complete the following checks:

(a) verify every citation and quotation against the original primary source to ensure the cited passage exists and supports the claim. (b) independently verify or reproduce every numerical, mathematical, or statistical result. (c) check every transcript segment against the original recorded audio. (d) test all generated software code against test cases specified independently by the researcher, rather than relying on test cases suggested by the AI tool. (e) confirm that all generated text reflects claims the researcher can independently explain and defend.

5.5 **The named-verifier rule.** Every research output must have one identified individual who is personally accountable for ensuring verification occurred. For a doctoral dissertation, that person is the doctoral student. For a co-authored paper, it is the corresponding author unless an explicit alternative division of responsibility is recorded in the authorship statement. For a grant proposal, it is the lead principal investigator.

5.6 **Documenting verification.** Verification must be documented rather than merely claimed. Acceptable documentation includes a citation-verification log, a version-controlled code repository showing manual test suites, a transcript audit note indicating which audio segments were checked, or recorded notes from an advising meeting confirming the checks performed.

5.7 The following core judgment tasks are strictly reserved for humans across all disciplines at [INSTITUTION NAME], with no exceptions permitted: examiner evaluation during an oral defense, an advisor's approval of a dissertation chapter, a peer-review assessment of a submitted manuscript, a determination of co-authorship credit, an institutional research-ethics approval decision, and any formal finding under Clause 14.

> ***Drafting note.*** *This clause is what moves an institution past a basic plagiarism policy, and it is usually the first section an external reviewer will evaluate. The Australian Research Council and National Health and Medical Research Council joint policy states that AI-generated content must be verified and should not replace expert opinion or judgement. It provides the clearest regulatory benchmark in the current literature (§2.1). However, stopping at a single general sentence achieves very little in practice. Clauses 5.4 and 5.6 turn that high-level requirement into a concrete checklist that leaves a verifiable audit trail. You should tailor Clause 5.6 to match your institution's actual record-keeping systems, because requiring documentation that nobody actually collects will cause the policy to be ignored within a year.*
>
> *Clause 5.7 is straightforward and absolute. Four of its six items are designated as strictly human tasks in the report's research task taxonomy (Appendix G): oral examination evaluation, supervisor chapter sign-off, authorship decisions, and peer-review recommendations.*
>
> ⚠ ***Jurisdictional variation.*** *Where an institution uses a formal legal declaration upon thesis submission (such as the German eidesstattliche Versicherung), you should embed this verification obligation directly into that declaration rather than creating a separate form. As the report notes, German universities successfully updated their existing statutory declarations to cover AI disclosure without rewriting their underlying legal frameworks (§1.3), which is the cleanest and most enforceable path where that tradition exists.*
>
> *Framework key: D1-policy, D1-process (the report §2.1, §4.2, Appendix G). Evidence of operation: the published Schedule A with a recorded version date, a sample of completed thesis verification logs, and an advising template that includes the required verification sign-offs.*

Clause 6 — Permitted and prohibited uses by research activity

6.1 Rules by AI-use mode

Mode	Permitted uses	Required conditions	Prohibited uses
Search	Scoping literature, identifying candidate sources, discovering authors, and finding relevant datasets	Every citation that remains in the work must be resolved to the primary source before inclusion (Clause 5.4(a)), with the tool name and access date recorded	Treating an unverified AI-generated bibliography as accurate, or citing a source the researcher has not personally read
Co-author	Editing, summarizing, adjusting tone, translating draft text written by the researcher, or generating initial drafts from the researcher's detailed outline	Full disclosure under Clause 7, and the researcher must be able to explain and defend every sentence independently	Generating substantive claims the researcher cannot defend, or presenting AI-paraphrased published literature as original human writing

Mode	Permitted uses	Required conditions	Prohibited uses
Validator	Running structured critiques, finding potential errors, and conducting devil's-advocate or pre-mortem reviews	The critique must be run through a model from a different frontier-model family than the one used to draft the argument, and model agreement must be treated as potential sycophancy rather than validation	Using an AI critique to substitute for required peer review, supervisory review, or expert committee evaluation
Tutor	Explaining unfamiliar concepts, working through technical examples, or practicing academic arguments	Corpus-grounded tools should be used whenever available, and the researcher must demonstrate independent understanding before AI assistance counts as learning	Using tutor outputs to bypass the core training milestones the degree program is designed to develop

Table 5. Permitted uses, required conditions, and prohibited activities across four research AI operational modes.

6.2 Rules by research activity

Schedule A sets the baseline classification for each research activity. The table below outlines the operative institutional rules:

Research activity	Class	Rule at [INSTITUTION NAME]
Initial literature scoping	Labor	Permitted. Citations must be resolved to primary sources before inclusion. Retrieval-grounded tools (Class 6) are strongly preferred over open web search (Class 1).
Citation checking during drafting	Labor	Permitted as a preliminary screening tool. The researcher must personally verify all citations.
Deciding which literature is foundational and resolving conflicting findings	Judgment	Permitted only for humans. The intellectual synthesis is the core contribution. AI tools may organize notes, but they cannot decide the argument.
Drafting text from a detailed researcher-generated outline	Labor	Permitted with appropriate disclosure under Clause 7.
Generating content the researcher cannot independently explain or defend	Judgment	Prohibited.
Editing language, refining style, or reducing word count on a researcher's own text	Labor	Permitted. Disclosure is governed by the threshold in Clause 7.2.
Translating a researcher's own original draft	Labor	Permitted, provided the original source text is retained and the researcher takes full ownership of the final translation. Institutions may choose tighter rules (see drafting note).
Generating code for data cleaning and statistical pipelines	Labor	Permitted. The researcher must review the code diffs, run manual tests, and take personal responsibility for the codebase.

Research activity	Class	Rule at [INSTITUTION NAME]
Selecting research methodologies and designing analytic plans	Judgment	AI tools may help brainstorm options, but the researcher must select the method and document the rationale.
Testing statistical assumptions and deciding on data exclusions	Judgment	Must remain a recorded human decision.
Developing hypotheses prior to data collection	Judgment	AI tools may suggest ideas, but the researcher must commit to the hypotheses in a pre-registered plan.
Establishing inclusion and exclusion criteria for systematic reviews	Judgment	The selection criteria define the research method and must be set by human researchers.
Machine-learning-assisted screening of titles and abstracts	Labor	Permitted. The researcher must personally review all borderline cases.
Transcribing interviews and focus groups	Labor	Permitted, provided original audio files are retained and audited against the text. Tool tier must follow Clause 8.
Creating illustrative diagrams, teaching graphics, or cover art	Labor	Permitted with appropriate disclosure.
Generating or altering empirical data figures (graphs, microscopy, gel images, structural models)	Prohibited	Strictly prohibited. AI tools cannot generate or modify data figures unless the AI system itself is the explicit object of study and this is documented in the methods.
Running adversarial reviews on one's own manuscript	Judgment (AI-assisted)	Permitted under the validator conditions in Clause 6.1. Does not fulfill human supervisory or peer-review obligations.
Examiner evaluation during an oral defense	Humans only	Strictly prohibited for AI systems (see Clause 11).
Advisor approval of a dissertation chapter	Humans only	Strictly prohibited for AI systems.
Determining co-authorship credit	Humans only	Strictly prohibited for AI systems. An AI system cannot be credited as an author.
Formulating peer-review recommendations	Humans only	Strictly prohibited for AI systems (see Clause 12.4).

Table 6. Institutional compliance rules and baseline classifications assigned to specific research activities.

6.3 Standing prohibitions

The following practices are strictly prohibited across all research disciplines and tool tiers:

(a) uploading research data into a tool tier that is not permitted under Clause 8. (b) uploading confidential manuscripts, draft dissertations, grant proposals, or unpublished datasets into any tool outside Tier B or Tier C. (c) presenting AI-generated material as the unaided work of a human, including in formal declarations. (d) using AI systems to invent, alter, or impute empirical research data presented as observed fact. (e) using the output of commercial AI-detection tools (Class 11) as disciplinary evidence in any proceeding under Clause 14. (f) using an AI system to make or determine decisions regarding academic admission, student progression, grading, degree completion, employment, or disciplinary action.

> ***Drafting note.*** *This clause incorporates two deliberate design choices. First, classifying a task as a labor task is not a blank check: every labor task still requires verification and transparency. Second, classifying an activity as a judgment task is not a total ban on software: it prohibits substituting AI for human decision-making, rather than forbidding researchers from using tools to explore options (the report, Appendix G). It is important to emphasize this distinction in training programs, because if researchers believe the policy simply bans tool use, they will continue using AI secretly and stop disclosing their methods.*
>
> *Two specific rows often generate debate during drafting. Regarding* ***Translation****, one university in the report's study banned AI translation of drafts written in another language on the grounds that translating raw text amounts to ghostwriting. If you have a significant multilingual graduate cohort, make this decision carefully. A complete ban places an unequal burden on researchers whose first language is not English, and Clause 6.3(e) specifically exists because commercial detection tools show documented bias against this exact group. Regarding* ***Data figures****, this policy's strict prohibition reflects broad consensus among academic publishers rather than a statutory mandate, and the report highlights this as a clear example of publisher rules leading institutional policy (Appendix C, Class 9).*
>
> *Framework key: D2-policy, D2-process (the report §2.2, §3.3, Appendix G). Evidence of operation: the task table published on the university research website, along with identical rules integrated into graduate curricula and faculty briefings.*

Clause 7 — Disclosure obligations

7.1 Formal disclosure is mandatory for all **substantive use** of AI in any research output, as defined in Clause 3.7. Disclosure is not required for routine incidental use.

7.2 Every disclosure statement must clearly record the following details: the tool name along with the vendor or development team, the specific model name and version (or the access date if no version is listed), the date or timeframe of use, the AI-use mode or modes involved, the specific research task where the tool was applied, the extent of the tool's contribution, and the name of the researcher who completed the verification required by Clause 5.4.

7.3 Disclosures must be documented in the specific locations designated for each type of research output:

Output type	Required disclosure location
Doctoral dissertation or thesis	The formal thesis declaration form and the front matter, as well as the methodology chapter where AI assisted data collection or analysis
Journal article or conference paper	The disclosure section required by the publisher, along with the institutional deposit record
Preprint	The methodology or acknowledgments section
Dataset or software repository	The repository README file and the project data-management plan
Grant application	The specific application section required by the funder, along with the internal pre-award institutional file

Output type	Required disclosure location
Ethics application	The formal application form, whenever AI touches participant recruitment, data processing, or participant interactions
Funder or industry partner report	The main body of the report, placed directly where the AI-assisted work appears

Table 7. Required disclosure locations for artificial intelligence use across different research output types.

7.4 The official institutional disclosure standards and statement templates are published in **P2: AI-Use Disclosure Standard and Statement Templates**. Researchers should use the standard P2 template unless a target publisher or funder mandates a specific alternative format. When an external format is required, researchers must complete that form and retain a copy of the P2 statement in their institutional project file.

7.5 If an external publisher or funder enforces a stricter or more granular disclosure standard than this policy, the external standard governs the submitted work while this policy governs the internal institutional record.

7.6 A researcher who discloses AI use openly and in good faith will not face adverse institutional action merely for having used an AI tool. Disciplinary consequences apply to prohibited uses, unverified work, or failures to disclose, rather than to disclosure itself.

7.7 Disclosure records must be preserved for [N] years following final publication or degree conferral, in alignment with the university's research-records retention schedule.

> ***Drafting note.*** *Clause 7.6 serves an important practical function. When the British Medical Journal introduced a structured AI disclosure field on April 8, 2024, it recorded an overall disclosure rate of just 5.7 percent across 25,114 submissions to 49 journals over the following seven months (the report §1.1, §3.3). When disclosure feels like a confession of wrongdoing, it reproduces that exact reporting gap. The policy should state clearly that disclosure is simply a transparent description of research methodology.*
>
> *The requirements in Clause 7.2 combine the most useful elements from leading publisher and university policies. You can simplify this list for your first year of implementation, but you should keep the tool name, model version, research step, and named verifier. Those four elements are what make a disclosure auditable.*
>
> *Framework key: D2-policy, D2-systems, D2-process (the report §2.2, §3.3). Evidence of operation: signed thesis declaration forms, the institutional disclosure registry, and audit samples comparing published papers against their internal deposit records.*

Clause 8 — Data classification and permitted tool tiers

8.1 What data a researcher may enter into an AI tool depends strictly on the security classification of that data, rather than on an individual's personal comfort with the software.

8.2 **Research data classifications.** For the purposes of this policy, [INSTITUTION NAME] classifies research data into four tiers:

Classification	Description	Typical examples
R1: Public	Information that is already lawfully in the public domain, with no restrictions on further sharing	Published journal articles, open datasets, and publicly available web text that researchers may lawfully process
R2: Internal	Non-public information where unauthorized disclosure would cause no significant harm	Teaching drafts, internal methodology notes, and routine administrative research records
R3: Confidential	Unpublished research material where disclosure could harm project integrity, external partners, or individuals	Unpublished manuscripts, working drafts under review, de-identified participant data, unawarded grant proposals, commercial research data, and dissertations prior to defense
R4: Restricted	Highly sensitive data subject to strict legal, ethical, or contractual controls	Identifiable human participant data, special-category medical data, Indigenous data governed by community agreements, export-controlled materials, and datasets with third-party processing restrictions

Table 8. Definitions, descriptions, and representative examples for institutional research data classifications.

8.3 The [ROLE: Data Protection Officer] and [ROLE: Chief Information Officer] jointly maintain the formal mapping between these research tiers and the university's overarching information-security framework. If the two classification schemes ever conflict, the stricter rule applies.

8.4 **Tool deployment tiers.**

Deployment tier	Definition
Tier A: Public	Commercial or free consumer tools used without an institutional agreement. The vendor may use user prompts to train models by default, and there are no institutional data-residency guarantees.
Tier B: Enterprise-tenanted	Commercial platforms covered by an institutional contract that explicitly disables model training on university inputs, provides enforceable data-residency terms, supports institutional audit logging, and includes a completed privacy impact assessment.
Tier C: Institution-controlled	Tools hosted directly on university computing infrastructure or running locally on an offline institutional device, such as on-premises clusters, self-hosted models, or open-weights models run locally.

Table 9. Definitions and operational characteristics of institutional artificial intelligence tool deployment tiers.

8.5 **Permitted combinations of data and tool tiers.**

Data classification	Tier A: Public	Tier B: Enterprise	Tier C: Institution-controlled
R1 Public	Permitted	Permitted	Permitted
R2 Internal	Prohibited	Permitted	Permitted
R3 Confidential	Prohibited	Permitted if the institutional contract and data-residency terms cover the data type	Permitted

Data classification	Tier A: Public	Tier B: Enterprise	Tier C: Institution-controlled
R4 Restricted	Prohibited	Prohibited unless explicitly approved in writing by [ROLE] following a formal impact assessment	Permitted if authorized by ethics approval, data-governance agreements, or contract terms

Table 10. Permitted combinations of research data classifications and artificial intelligence tool deployment tiers.

8.6 Where a research contract, ethics approval, data-transfer agreement, or Indigenous data-governance protocol imposes tighter restrictions than this matrix, those specific terms take precedence.

8.7 Interaction histories and session logs generated within Tier B and Tier C systems are treated as official institutional records. They must be retained for [N] years and may be accessed during research-integrity investigations, formal complaints, or supervisory reviews under the safeguards established in Clause 14.6.

> ***Drafting note.*** *You should link this policy directly to your university's existing data-classification system (Clause 8.3) rather than creating a separate framework from scratch, because your IT security team will already have standards in place. There are two common failure modes here. First, if Tier B enterprise tools are slow or difficult to access, this raises the risk that researchers will use consumer Tier A tools with confidential R3 data and simply avoid telling you. The classic example in industry is Samsung's experience in March 2023, where three separate leaks of confidential source code occurred within twenty days, leading to an outright ban on consumer generative AI tools (the report §2.3). You need to provide viable tools before you enforce restrictions. Second, setting an overly rigid rule for R4 data can block the very projects that need modern tools most, which is why Clause 8.5 provides a clear written approval pathway rather than an absolute ban.*
>
> ⚠ ***Jurisdictional variation.*** *Data residency and cross-border transfer rules differ significantly across regions. In the European Economic Area and the United Kingdom, you must state data residency requirements explicitly in your Tier B contracts. Australian universities should align this section with Privacy Act requirements under Clause 9.7 as well as national Indigenous data-sovereignty principles. Canadian institutions must comply with the federal Personal Information Protection and Electronic Documents Act (PIPEDA) and provincial private-sector privacy statutes (such as Quebec Law 25, Alberta PIPA, and British Columbia PIPA). In mainland China, internal academic research sits outside the Cyberspace Administration's public algorithm-filing rules, but any public-facing tool deployment falls under them. Finally, United States export-controlled materials (such as ITAR or EAR data) must be classified as R4 in all cases.*
>
> *Framework key: D3-policy, D3-systems (the report §2.3). Evidence of operation: the published classification matrix, enterprise tenancy contracts showing model training disabled, and data-retention schedules covering AI interaction logs.*

Clause 9 — Procurement and approved tooling

9.1 [INSTITUTION NAME] maintains an official register of approved AI research tools in Schedule B. For every tool, the register records the tool class (Clause 3.4), the security tier (Clause 8.4), permitted data classifications, the formal approval decision and date, the exact model version evaluated, test results across the six criteria in Clause 9.2, the scheduled re-evaluation date, and the designated system owner.

9.2 **The procurement evaluation.** Before approving any AI tool for research use, [ROLE] evaluates the tool against six observable technical criteria and records the evidence. Marketing statements from vendors do not count as proof:

(a) **Verifiable citations.** The tool must produce checkable citations for factual statements, and the evaluation team must manually verify a test sample of [N] citations during procurement to record the accuracy rate. (b) **Data residency and training protections.** The vendor contract must explicitly prevent user inputs from being used for model training, and must clearly identify where data is processed and stored. (c) **Uncertainty reporting.** The tool must communicate confidence scores or uncertainty metrics directly, or the register must document that it does not and outline the resulting restrictions on use. (d) **Version stability and reproducibility.** The specific model version must be identifiable and lockable, and the tool's behavior under identical prompts must be documented, including any non-deterministic variation. (e) **Audit capabilities.** Interaction logs must be captured and retrievable by the institution for the retention period specified in Clause 8.7. (f) **Evaluation of open and local alternatives.** Where an open-weights or locally hosted tool could handle the research workload, the register must record whether that option was evaluated and why the commercial tool was chosen.

9.3 **The manual citation verification test in 9.2(a) is mandatory and cannot be satisfied by vendor marketing.** A 2024 study by Stanford RegLab evaluating three retrieval-augmented legal research tools discovered hallucination rates between 17 and 33 percent, despite vendor claims that retrieval mechanisms had solved citation accuracy (the report §2.3). In fact, retrieval grounding often shifts the failure pattern from obvious fictional citations to subtle misattributions of real sources, which are much harder for human readers to detect. The evaluation team must record the observed pass rate in the register and share it across the institution.

9.4 **Re-testing schedule.** Every approved tool must be re-evaluated against Clause 9.2 at least once every [12] months, and immediately whenever a provider updates the default model version. A tool that misses its re-evaluation date moves to provisional status and cannot be used with R3 or R4 data until testing is complete.

9.5 **Unapproved shadow tooling.** Researchers must not process R2, R3, or R4 research data using tools that are not listed on the approved register. The [ROLE] maintains an expedited review process to evaluate new tool requests within [10] business days.

9.6 **Exception pathway.** If a specialized research project requires a tool that cannot meet all criteria in Clause 9.2, [ROLE] may grant a formal written exception. The exception must record the specific project needs, the unmet technical criteria, the compensating safety controls, the permitted data classifications, and an expiration date. All exceptions must be reported to [COMMITTEE] at each scheduled meeting.

9.7 **High-risk deployment surfaces.** [INSTITUTION NAME] will not deploy or approve AI systems that make or materially inform decisions regarding admissions, student grading, academic standing, or exam proctoring until the required statutory impact assessment is completed, registered where legally mandated, and approved by [ROLE].

> ***Drafting note.*** *Clause 9.2 turns the six technical criteria in the report into a concrete procurement gate, but it only works if you enforce Clause 9.3. Accepting a vendor's claim that their system is "hallucination-free" leaves you with a policy that looks good on paper but fails in practice. You should size the verification sample in Clause 9.2(a) so that it produces meaningful data without overwhelming staff. Having a research librarian check twenty to fifty sample citations is a very workable approach that yields a clear accuracy figure for your governance committees.*
>
> *In Clause 9.4, the trigger for version changes is more important than the annual calendar review. AI vendors frequently change underlying model versions without advance notice, and a system evaluated in March may give completely different answers to the same research prompt by September. Clause 9.5 is necessary because having a slow approval gate simply encourages unmonitored shadow IT, which creates far more institutional risk than a responsive evaluation process.*
>
> ⚠ ***Jurisdictional variation: Clause 9.7 carries this policy's two most critical compliance dates.*** *In the* ***European Union****, the EU AI Act classifies AI systems used for university admissions, learning assessments, academic placement, and exam proctoring as high-risk systems (Annex III §3), and the Article 27 Fundamental Rights Impact Assessment becomes mandatory for these uses on* ***August 2, 2026*** *(the report §2.3, §2.5, §4.2, Appendix A). In* ***Australia****, amendments to the Privacy Act 1988 require an organization's privacy policy to disclose the types of personal data used in automated decision-making and outline the kinds of decisions made, taking effect on* ***December 10, 2026*** *(the report §2.5, §4.2). You should align Clause 9.7 with the Data Protection Impact Assessment requirements in the United Kingdom, relevant state and federal grant compliance rules in the United States, or federal and provincial privacy laws in Canada. Keep only the legal references that apply to your institution and remove the rest.*
>
> *Framework key: D3-policy, D3-systems, D3-process, D5-policy (the report §2.3, §2.5). Evidence of operation: the completed Schedule B tool register with detailed test scores, dated re-evaluation records, and signed impact assessments with their official filing numbers.*

Clause 10 — Supervision and research-training obligations

10.1 **The supervisor–student agreement.** Every doctoral student and their advisory team must complete a formal written AI-use agreement (**G1: Supervisor–Student AI-Use Agreement**) within [three] months of starting their program. This agreement records what AI use is expected, permitted, or restricted within the specific research project, and establishes how verification will be documented. The agreement must be reviewed and updated at every formal milestone, and whenever the project's methodology changes significantly.

10.2 **Core curriculum learning outcomes.** The graduate research-training program at [INSTITUTION NAME] must include the six essential AI-era competencies as explicit learning outcomes, complementing rather than replacing existing coursework on research integrity and validity:

(a) primary source citation verification. (b) model and parameter specification. (c) prompt design and decision-branching discipline. (d) model heterogeneity in adversarial testing. (e) sycophancy detection and human verification discipline. (f) structured failure-mode documentation.

10.3 **Advisor and supervisor training.** Advisors and supervisors must complete dedicated training on this policy and the core competencies in Clause 10.2 before taking on new graduate research students, and must complete a refresher module at least once every [24] months. Training completion is recorded by [ROLE] and reported annually to [COMMITTEE].

10.4 **Examiner briefing.** Internal and external examiners must receive a briefing on Clause 11 prior to formal appointment. If an external examiner cannot complete institutional training, the university must include the Clause 11 rules in the official appointment letter and obtain written confirmation of agreement.

10.5 **Research induction.** University induction programs for new graduate research students and incoming academic staff must cover this policy, Schedule A, and the disclosure standards in P2.

10.6 **Supervisory technical competence.** Advisors and supervisors must not approve any dissertation chapter, data analysis, or manuscript draft containing AI-assisted elements that they cannot independently evaluate. If an advisor lacks the technical expertise to evaluate a specific AI method, they must arrange for a qualified colleague to review the work and record that this review took place.

10.7 Records of completed training, signed G1 agreements, and milestone review dates must be maintained in the graduate student management system and be auditable by school or faculty.

> ***Drafting note.*** *Clause 10.3 is often the most challenging requirement to implement across the entire policy, and it is the section most likely to face pushback during committee review. However, the finding from the report is very direct: doctoral students cannot build competencies that their advisors do not understand or model (§2.4). If you remove this requirement, the policy becomes a one-sided document that holds students accountable while exempting the faculty who guide them.*
>
> *Clause 10.6 requires an advisor who cannot evaluate an AI-assisted analysis to bring in a colleague who can, which is how a graduate school naturally identifies where its supervisory skill gaps are located. Two Class D institutions in Europe have codified requirements like this: KU Leuven's structured training path for doctoral researchers and the University of Helsinki's literacy requirement tied to institutional courses (the report §3.2, §4.2). I recommend reviewing both models when drafting your local version of Clause 10.2.*
>
> *Framework key: D4-policy, D4-people, D4-systems, D4-process (the report §2.4, §5.3). Evidence of operation: signed G1 agreements sampled across departments, curriculum documents listing the six learning outcomes, and supervisor training records with completion dates.*

Clause 11 — Examination and the thesis

11.1 **Thesis declaration.** Every thesis or dissertation submitted for examination must include a signed declaration detailing all AI use in accordance with Clause 7. The declaration must list the tools used, the operational modes, the specific research stages involved, and a formal statement confirming that all verification requirements under Clause 5.4 were satisfied.

11.2 **Oral defense and defensibility.** A doctoral student must be able to describe and defend all uses of AI in their research, as well as the complete intellectual content of the thesis, during the final oral examination or defense. An inability to explain or defend the work is evaluated as an academic examination matter, rather than automatically triggering a misconduct proceeding under Clause 14.

11.3 **Prior supervisory awareness.** Substantive AI use in dissertation research or writing must be discussed with the supervisory team as it occurs and recorded in the G1 agreement. A doctoral student must never disclose substantive AI use for the first time upon final submission.

11.4 **Examiner confidentiality.** Examiners must not upload any portion of a draft dissertation, submitted thesis, or examination report into any generative AI tool, including an enterprise-tenanted or locally executed system. A thesis under examination is confidential R3 data, and examination materials remain outside the AI-tool workflow.

11.5 **Prohibition on commercial detection tools.** Examiners must not use commercial AI-detection software to evaluate a thesis or justify an examination recommendation, and must not cite detection scores as evidence. High false-positive rates and documented biases against multilingual writers make these tools unfit for academic assessment (the report, Appendix C, Class 11).

11.6 **Examiner judgment.** All examination recommendations and evaluations must be made personally by the examiners. No part of the evaluative judgment may be delegated to an AI tool.

11.7 **Handling suspected undisclosed use.** An examiner who suspects undisclosed or unverified AI use should document the specific passages and concerns in writing and refer the matter to [ROLE]. The examiner must not attempt to adjudicate the issue independently or raise unverified accusations during the oral examination. The defense proceeds or is paused according to the established university [PROCEDURE].

11.8 Where [INSTITUTION NAME] uses a formal statutory declaration upon thesis submission, the AI disclosure required under Clause 11.1 must be integrated directly into that declaration, avoiding the creation of duplicate forms.

> ***Drafting note.*** *Clauses 11.4 and 11.5 draw directly on the two UK Class D policies that address examiners explicitly: University College London Doctoral School's guidance on transparency and King's College London's framework for students, supervisors, and examiners. The report notes that both institutions addressed examiner responsibilities and defense protocols far more thoroughly than most peer universities (§3.2).*
>
> *Clause 11.2 adapts the University of Toronto School of Graduate Studies model, which requires students to explain and defend any AI use alongside their thesis findings during the oral defense. Toronto is classified at* ***Class C****, and the report highlights this clause as an unusually clear and practical rule. It shifts the institutional question away from a punitive "did you use software?" toward an academic "can you defend your scholarship?", which is the exact question an examination committee is qualified and expected to evaluate.*
>
> *Clause 11.7 is critical for due process. When an examiner raises unverified allegations of AI misuse inside an oral defense, it turns an academic assessment into an unfair disciplinary hearing. Routing concerns through a separate review process protects both the student and the integrity of the examination. Be sure to check this language against your university's current academic appeal regulations.*
>
> ⚠ ***Jurisdictional variation.*** *Examination formats vary widely across systems, ranging from public defenses before large committees to closed oral vivas with two external examiners, or entirely written assessments without an oral component. Clause 11.2 assumes an oral defense. Where degrees are evaluated purely on written submissions, the equivalent requirement is that the student must provide written explanations of their AI methods upon request from the examiners, which places greater weight on the initial submission declaration in Clause 11.1.*
>
> *Framework key: D1-process, D2-process, D4-process (the report §2.1, §4.2, §5.3). Evidence of operation: signed thesis declaration forms, examiner appointment letters containing Clause 11 terms, and records of formal referrals under Clause 11.7.*

Clause 12 — Publication, preprints, peer review, and grant applications

12.1 **Authorship standards.** An AI system cannot be listed as an author or co-author on any research output. Authorship requires personal accountability, which a software tool cannot assume. This position is held consistently across all eighteen major academic publishers surveyed in the report (§3.3).

12.2 **Publisher disclosures.** Researchers must comply with the specific AI disclosure guidelines of their target publication, and must archive a copy of the institutional disclosure statement under Clause 7.4. Where a publisher does not provide a specific format, researchers must use the standard P2 template.

12.3 **Figures and empirical data.** Clause 6.2 applies to all submissions. While illustrative diagrams or graphical abstracts created with AI assistance are permitted with disclosure, generating or modifying empirical data figures with AI tools is strictly prohibited.

12.4 **Peer-review responsibilities.** Any university researcher serving as a peer reviewer for a journal, conference, granting agency, or external institution must:

(a) never upload a confidential manuscript, grant proposal, or working draft into any tool outside Tier B or Tier C, because material under review is confidential R3 data. (b) comply fully with the specific AI policy of the journal or funding agency, including total bans where enforced. (c) formulate all evaluative recommendations and review comments personally. (d) disclose any permitted AI assistance to the editor, and to the authors when required by the venue.

12.5 **Grant applications.** Research proposals submitted by [INSTITUTION NAME] must represent the original intellectual work of the named investigators. Proposals or proposal sections substantially generated by AI tools must not be submitted. This reflects the strictest applicant standard in the report's funder cohort: the United States National Institutes of Health originality rule (NOT-OD-25-132, §3.1). [INSTITUTION NAME] adopts this rule as its baseline standard because a funder's originality requirement applies to all applicants regardless of their home institution's location.

12.6 **Funder requirements register.** The [ROLE: Research Office] maintains an up-to-date summary of AI guidelines across major research funding bodies in Schedule D. This register is reviewed at least [annually], and updates are communicated directly to research leaders.

12.7 **Research ethics applications.** Whenever AI systems are used in participant recruitment, direct participant interactions, the analysis of identifiable personal data, or any process that affects human subjects, the details must be explicitly disclosed in the research-ethics application and reviewed by [COMMITTEE: Human Research Ethics].

> ***Drafting note.*** *Clause 12.5 represents a deliberate institutional simplification. Funder rules range from having no disclosure requirements at all (the Swedish Research Council is the most flexible in the report's sample) to strict originality requirements like those of the US National Institutes of Health. If an institutional policy simply tells researchers to check individual funder guidelines on their own, compliance is usually inconsistent. Adopting the strictest funder standard as the university default costs very little, because no research funder encourages AI-generated grant applications, and having a clear baseline prevents avoidable compliance errors. If leadership prefers tracking rules on a funder-by-funder basis, you can replace Clause 12.5 with a direct cross-reference to Schedule D.*
>
> *Clause 12.4 is vital because peer review is an area where an individual researcher can inadvertently breach confidentiality with a single paste into a public browser window. Sixteen of the eighteen publishers in the report's study already explicitly prohibit uploading manuscripts under review into public tools. While Cambridge University Press and SAGE do not explicitly address this in their primary policy summaries, that omission should never be interpreted as permission to upload confidential submissions (§3.3).*
>
> *Framework key: D2-policy, D2-process, D3-policy (the report §2.2, §3.1, §3.3). Evidence of operation: the Schedule D funder register with recent review dates, institutional grant submission checklists incorporating Clause 12.5, and research-ethics forms with dedicated AI disclosure fields.*

Clause 13 — Roles and responsibilities

13.1 Institutional accountability under this policy is distributed across four core operational axes. Every axis must have a designated individual leader:

Axis	Designated leader	Scope of accountability	Primary owned artifact
Policy	[ROLE (recommended: Deputy Vice-Chancellor (Research), Vice-President or Vice-Provost for Research, or equivalent)]	Maintaining policy currency, scope, institutional alignment, and formal committee approvals	The official policy text and published version history
People	[ROLE: Dean of Graduate Studies] together with [ROLE: Director, Researcher Development]	Ensuring training provision, tracking completion rates, and building AI competencies among students, supervisors, and examiners	Course learning outcomes, training completion logs, and examiner briefing materials
Systems	[ROLE: Chief Information Officer] together with [ROLE: Data Protection Officer] and [ROLE: University Librarian]	Managing the approved tool register, data tiering, log retention, and required impact assessments	Schedule B, data-classification mappings, and completed impact-assessment files
Process	[ROLE: Chair, Graduate Studies Committee] together with [ROLE: Director, Research Integrity]	Overseeing advising workflows, progression milestones, examination protocols, and breach responses	Milestone procedures, examination forms, and Clause 14 investigation workflows

Table 11. Designated leaders, accountability scopes, and primary artifacts across four institutional governance axes.

13.2 **Ownership matrix across dimensions and axes.** [INSTITUTION NAME] assigns an operational leader for each of the twenty cells in the benchmarking framework. The table below is completed upon initial policy adoption and republished following each review. It serves as the baseline input for institutional self-assessments using instrument A1:

Maturity dimension	Policy	People	Systems	Process
D1: Human-in-the-loop discipline	[ROLE]	[ROLE]	[ROLE]	[ROLE]
D2: Responsible use in practice	[ROLE]	[ROLE]	[ROLE]	[ROLE]
D3: Tooling that promotes responsible use	[ROLE]	[ROLE]	[ROLE]	[ROLE]
D4: AI-literate humans	[ROLE]	[ROLE]	[ROLE]	[ROLE]
D5: Institutional benchmarking grid	[ROLE]	[ROLE]	[ROLE]	[ROLE]

Table 12. Operational leadership assignments across maturity dimensions and institutional governance axes.

13.3 **Individual responsibilities.**

(a) *Researchers and graduate students* must comply with this policy, verify all AI outputs before incorporating them, disclose substantive use, and adhere strictly to permitted data and tool tiers. (b) *Advisors and supervisors* must complete the G1 agreement with their students, model core AI competencies, review student AI use at every progression milestone, and satisfy the technical competence requirements in Clause 10.6. (c) *Examiners* must follow the conduct and confidentiality rules in Clause 11. (d) *Department chairs, heads of school, and associate deans for research* must maintain discipline-specific extensions to Schedule A and identify emerging tools used in their units for registration. (e) *[COMMITTEE]* receives the annual self-assessment report under Clause 15.4 and reviews exceptions submitted under Clause 9.6.

> ***Drafting note.*** *Table 13.2 is the most critical governance tool in this entire document for any team planning to evaluate institutional maturity. The report requires all four axes to operate for the established level. The pack applies that rule by setting each dimension at the level of its lowest-scoring axis. It calls that limiting axis the binding axis. This means that leaving an unassigned owner in any cell holds that entire dimension back. In other words, an institution cannot be established on human-in-the-loop discipline if nobody has been assigned to manage the systems that support it. Completing this matrix is a two-hour administrative task that quickly highlights real operational gaps. The most common outcome on a first pass is discovering that the policy axis has clear leadership while the other three axes have no designated owners. That exact pattern is how the report defines a nascent maturity level on Dimension 5: an institution with formal policies on paper, but where only the policy axis has an assigned owner (Appendix A).*
>
> *Be sure not to combine roles so much that individual accountability is lost, and avoid assigning oversight to broad committees where an individual role should be named. You cannot ask a committee why a specific operational task was missed.*
>
> *Framework key: D1-people, D2-people, D3-people, D4-people, D5-people, D5-policy (the report §2.5, §5.1). Evidence of operation: the completed 13.2 table published alongside the policy, committee minutes approving the ownership list, and designated staff responding clearly when asked who manages D3-systems.*

Clause 14 — Breach and proportionate response

14.1 **Classification of cases.** Any suspected policy violation must be classified into one of four distinct categories before deciding on an institutional response. Each category requires different evidence and distinct remedial actions:

Category	Definition	Required evidence	Initial institutional response
(a) Undisclosed use	Substantive AI use that should have been disclosed was omitted, but the underlying research work is technically sound	The research output, the disclosure statement (or its absence), and the researcher's explanation	Correcting the record by adding a disclosure statement, accompanied by an advising discussion and, for graduate students, a supervisory note
(b) Disclosed but unverified use	AI use was openly disclosed, but the verification required by Clause 5.4 was not conducted	The research output, the disclosure record, the absence of verification logs, and the specific unverified claims	Correcting the work, requiring re-verification of the output by the researcher, assigning targeted training, and escalating if the work has already been published

Category	Definition	Required evidence	Initial institutional response
(c) AI-fabricated content in the record	Fabricated citations, invented quotations, false data, or altered images have entered the research record	The specific fabricated elements, documented alongside original primary sources showing the claims are unsupported	Referral to the research-integrity office, followed by formal correction, retraction, or withdrawal as appropriate, along with notifications to publishers, funders, or examiners
(d) AI-assisted research misconduct	Deliberate use of AI tools to falsify data, fabricate findings, plagiarize scholarship, or actively conceal unauthorized use	The formal investigative findings from the research-integrity office	Initiating formal research-misconduct proceedings under the institutional [POLICY]

Table 13. Violation categories, definitions, required evidence, and initial institutional responses for policy breaches.

14.2 **Proportionality.** Institutional responses must be proportionate to the actual harm done to the research record, taking into account the researcher's academic stage and prior history. A first-instance failure to disclose by a first-year graduate student is an advising and educational matter, rather than an academic misconduct proceeding.

14.3 **Accurate case classification is mandatory prior to action.** Treating a fabrication case (Category c) merely as traditional plagiarism understates the problem, because the primary harm is the corruption of the scientific record rather than simple misattribution. Conversely, treating a minor disclosure omission (Category a) as formal research misconduct over-responds, which quickly discourages open disclosure across the entire student community.

14.4 **Evidentiary standards.** All findings must be based on verifiable evidence in the research record: the output itself, primary sources, disclosure forms, verification notes, and the researcher's account. The output of commercial AI-detection tools is not evidence and cannot be cited in findings, investigative referrals, or supervisory notes (Clause 6.3(e)).

14.5 **Burden of proof.** The university carries the burden of demonstrating that a policy violation occurred. A researcher is never required to prove the negative claim that they did not use AI tools.

14.6 **Accessing user session logs.** Interaction logs stored under Clause 8.7 may be accessed during an inquiry only with the written authorization of [ROLE], based on a defined scope of review, and with advance notice to the researcher unless notifying them would compromise an active investigation into Category (d) misconduct.

14.7 **Good-faith disclosures.** Good-faith disclosure of AI use is not, by itself, grounds for adverse action. Any disclosed conduct remains subject to Clauses 5, 6, 8, and 14.

14.8 Disciplinary appeals are handled according to standard university [PROCEDURE].

> ***Drafting note.*** *The four categories in Clause 14.1 reflect the adjudication framework developed in the companion report (§5.4). Institutional committees that lack this four-part distinction tend to make two common mistakes: they either apply basic plagiarism rules to serious data fabrications (which trivializes the damage), or they launch formal misconduct investigations against graduate students who simply forgot a disclosure check (which frightens the entire cohort and stops people from being transparent).*
>
> *Clause 14.4 will almost certainly be challenged by someone on your committee who wants to run student papers through automated detection software. You should hold the line here. A 2024 study by Perkins and colleagues evaluating seven major AI detectors found an average baseline accuracy of just 39.5 percent, falling to 17.4 percent when simple text-editing techniques were applied, along with persistent error rates against non-native English writers (the report, Appendix C, Class 11). Allowing automated detection scores into disciplinary hearings introduces an unacceptably high error rate that falls disproportionately on international students.*
>
> ⚠ ***Jurisdictional variation.*** *National research-integrity frameworks establish the formal procedures for Category (c) and Category (d) matters, including required standards of proof, investigative processes, and mandatory external reporting to funding bodies. You should tie Clauses 14.1(c) and 14.1(d) directly into your national research-integrity code and your existing university misconduct processes, rather than creating an entirely separate disciplinary structure within this policy.*
>
> *Framework key: D2-process, D5-process (the report §5.4). Evidence of operation: an institutional case log categorized by violation type, records of formal appeals, and written operating procedures identifying the decision-maker for each category.*

Clause 15 — Review, cadence, and named owner

15.1 **Designated policy owner.** This policy is owned and maintained by [ROLE]. The policy owner is accountable for keeping the text current, delivering the annual self-assessment report required by Clause 15.4, and maintaining the completed ownership matrix in Clause 13.2.

15.2 **Review schedule.** The main policy text is reviewed at least once every [24] months. Schedules A, B, and D are reviewed at least once every [12] months. In addition, the approved tool register is updated whenever a scheduled re-test under Clause 9.4 changes a tool's approval status.

15.3 **Event-triggered reviews.** A formal review of this policy is triggered immediately, outside the regular calendar schedule, by any of the following occurrences: a new statutory or regulatory requirement in any jurisdiction where [INSTITUTION NAME] operates, a significant technical change in an approved tool or tool class, a Category (c) or Category (d) finding under Clause 14, or a major change in disclosure policies across the primary journals or funding agencies that support the university's research.

15.4 **Self-assessment and reporting.** The [ROLE] assesses [INSTITUTION NAME] against the five maturity dimensions and four operational axes at least [annually] using instrument A1, and presents a formal report to [COMMITTEE]. This report outlines the maturity level achieved on each dimension, identifies the specific limiting axis holding back each dimension, highlights the overall

institutional bottleneck, and details committed improvements for the upcoming year. The report does not compress the findings into a single institutional score, because the published findings in the companion report show that the true diagnostic value of the framework lies in the multidimensional pattern across all five areas (§4.3).

15.5 **Public reporting and benchmarking.** [INSTITUTION NAME] publishes its self-assessment scores and evaluation methodology openly. Publishing self-assessment results, incorporating external peer benchmarking, and documenting regulatory readiness plans are what distinguish the *leading* level from the *established* level on Dimension 5 (the report §4.2, Appendix A). As noted in the report, no university in the mid-2026 sample had yet published its full self-assessment scores against this framework because it had only recently been released, making the *leading* tier a forward-looking standard that early adopters can choose to fulfill (§4.2).

15.6 **Version management.** Every policy version must record its approval date, effective date, a summary of modifications, and the specific review trigger that prompted the revision. Policy changes do not retrospectively reclassify conduct permitted when undertaken, subject to continuing legal, ethical, contractual, confidentiality, funder, and publisher obligations. Superseded versions must be archived and remain accessible for [N] years.

> ***Drafting note.*** *The two-track review cadence in Clause 15.2 is an intentional design choice. The AI technical landscape evolves much faster than standard academic governance, and tying the software tool register to a two-year committee cycle means you will constantly be governing systems that no longer behave as they did when evaluated. Keep your schedules on a short annual cycle while keeping core policy principles on a stable two-year cycle.*
>
> *Clause 15.4 ensures this policy can be measured and audited in daily practice rather than remaining an aspirational document. Clause 15.5 is often debated vigorously in committee, because publishing an institutional assessment that shows lower scores in certain cells can feel uncomfortable. However, based on the research in the report, being transparent about areas that need development is precisely what separates a leading institution from an established one on Dimension 5.*
>
> *Framework key: D5-policy, D5-people, D5-systems, D5-process (the report §2.5, §4.2, §5.1). Evidence of operation: a dated policy version history, annual self-assessment reports recorded in [COMMITTEE] minutes, and publicly available benchmark scores.*

Schedules

This policy is not operational without these four supporting schedules. Each schedule is actively maintained by its designated owner and must display an updated version date.

Schedule A: Labor-and-judgment task demarcation. The institutional research task list, adapted from the taxonomy in Appendix G of the report and customized by each school or faculty for its specific disciplines. Fields: research task name, baseline classification (labor, judgment, judgment-humans-only, prohibited), required conditions, department extensions, and version date. *Owner: [ROLE: Director, Research Integrity].*

Schedule B: Approved tool register. The official list of evaluated software tools as defined in Clause 9.1. Fields: tool name, tool class (1–11), security tier (A, B, or C), permitted data classifications (R1–R4), approval decision and date, tested model version, evaluation records against the six

technical criteria (including measured citation accuracy), re-test due date, system owner, and current status. *Owner: [ROLE: Chief Information Officer].*

Schedule C: Data classification mapping. The formal mapping connecting research data tiers (R1–R4) to the university's general information-security classifications, including the permitted deployment matrix from Clause 8.5 and any specific overrides required by contracts, ethics approvals, or data-governance agreements. *Owner: [ROLE: Data Protection Officer].*

Schedule D: External obligations register. The register of external AI policies, including funder requirements, publisher disclosure standards for key target journals, and statutory mandates across each jurisdiction where [INSTITUTION NAME] operates, along with assigned internal owners, effective dates, and implementation status. *Owner: [ROLE: Director, Research Office].*

Local decisions this template leaves open

The drafting team should work through these eighteen institutional choices before submitting the final policy to [COMMITTEE] for approval. Each item represents a deliberate local governance decision rather than an oversight in the template:

#	Local decision	Relevant clause	Status
1	Choosing between Class C and Class D scope: deciding whether Clauses 10.2–10.4 and Clauses 11.4–11.7 are included or omitted	How to use this template	☐ Yes ☐ Partial ☐ No
2	Deciding whether individual schools and faculties may establish stricter local rules, or looser ones	2.4	☐ Yes ☐ Partial ☐ No
3	Setting the precise boundary between incidental use and substantive use for a multilingual student cohort	3.7, 6.2	☐ Yes ☐ Partial ☐ No
4	Deciding whether AI translation of a researcher's own drafts is permitted, restricted, or prohibited	6.2	☐ Yes ☐ Partial ☐ No
5	Determining the specific verification documentation the university will collect and audit in practice	5.6	☐ Yes ☐ Partial ☐ No
6	Setting the formal retention period for disclosure records and AI interaction logs	7.7, 8.7	☐ Yes ☐ Partial ☐ No
7	Mapping research data tiers (R1–R4) onto the existing institutional data-classification framework	8.3	☐ Yes ☐ Partial ☐ No
8	Sizing the test sample and setting the minimum accuracy threshold for manual citation verification	9.2(a), 9.3	☐ Yes ☐ Partial ☐ No
9	Establishing an institutional turnaround target for evaluating new tool requests	9.5	☐ Yes ☐ Partial ☐ No

#	Local decision	Relevant clause	Status
10	Identifying the specific regional or national regulatory compliance dates to cite in the policy text	9.7	☐ Yes ☐ Partial ☐ No
11	Deciding whether supervisor training is mandatory before a faculty member takes on a new graduate student	10.3	☐ Yes ☐ Partial ☐ No
12	Establishing the mechanism for communicating Clause 11 rules to external examiners who cannot complete university training	10.4, 11.4–11.6	☐ Yes ☐ Partial ☐ No
13	Deciding whether AI disclosure is integrated directly into an existing statutory submission declaration	11.8	☐ Yes ☐ Partial ☐ No
14	Establishing the institutional default for grant proposals: applying the strictest funder rule or tracking policies on a funder-by-funder basis	12.5, 12.6	☐ Yes ☐ Partial ☐ No
15	Assigning an accountable owning role to each of the twenty cells in Clause 13.2, allowing one person to hold more than one assignment	13.2	☐ Yes ☐ Partial ☐ No
16	Mapping each Clause 14 case category into existing university disciplinary and integrity procedures	14.1	☐ Yes ☐ Partial ☐ No
17	Establishing review cadences for the main policy text and its technical schedules	15.2	☐ Yes ☐ Partial ☐ No
18	Deciding whether the university will publish its institutional self-assessment scores openly	15.5	☐ Yes ☐ Partial ☐ No

Table 14. Local governance decisions, associated policy clauses, and adoption status for institutional drafting teams.

Crosswalk — clause to report section

Clause	Framework cells served	Related report section
1 — Purpose	D1-policy, D2-policy	§1.2 (the operational definition of responsible use), and §1.3 (the four policy classes)
2 — Scope and application	D1-policy, D5-policy	§2.5 (the four operational axes), and §2.3 (the regulatory landscape)
3 — Definitions	D2-policy, D3-policy	§2.2 (the four AI-use modes), §2.3, and Appendix C (the eleven tool classes)
4 — Principles	D1-policy, D2-policy, D3-policy, D4-policy, D5-policy	§2.1 through §2.5

Clause	Framework cells served	Related report section
5 — Human-in-the-loop obligation	D1-policy, D1-process	§2.1, §4.2 (Dimension 1 at the *leading* tier), and Appendix G
6 — Permitted and prohibited uses	D2-policy, D2-process	§2.2, §3.3 (publisher consensus on graphics and authorship), and Appendix G
7 — Disclosure obligations	D2-policy, D2-systems, D2-process	§2.2, along with §1.1 and §3.3 (disclosure practices and the reporting gap)
8 — Data classification and tool tiers	D3-policy, D3-systems	§2.3 (data residency and model-training controls)
9 — Procurement and approved tooling	D3-policy, D3-systems, D3-process, D5-policy	§2.3 (the six observable properties), §2.5 (regulatory requirements), and Appendix A row 3
10 — Supervision and research training	D4-policy, D4-people, D4-systems, D4-process	§2.4 (the six core competencies), and §5.3 (graduate school implementation)
11 — Examination and the thesis	D1-process, D2-process, D4-process	§2.1, §3.2 (examiner rules in Class D policies), §4.2, and Appendix C Class 11
12 — Publication, peer review, grants	D2-policy, D2-process, D3-policy	§3.1 (funder requirements), §3.3 (publisher standards), and §2.2
13 — Roles and responsibilities	D1-people through D5-people, D5-policy	§2.5 and §4.1 (all four axes required for *established*, with this pack using the lowest axis level), and §5.1
14 — Breach and proportionate response	D2-process, D5-process	§5.4 (the four violation categories), and Appendix C Class 11 (limits of detection tools)
15 — Review, cadence, owner	D5-policy, D5-people, D5-systems, D5-process	§2.5, §4.2 (Dimension 5 at the *leading* tier), §4.3 (avoiding single scores), and §5.1
Schedules A–D	D1-policy, D3-systems, D5-policy, D5-systems	Appendix G (Schedule A), §2.3 (Schedules B and C), and §2.5 and §3.1 (Schedule D)
Local decisions table	D5-policy, D5-process	§4.1 (non-prescriptive maturity targets), and §4.3

Table 15. Mapping of policy clauses to served framework cells and corresponding report sections.

instats

Michael J. Zyphur, PhD · Professor and Director, Instats · instats.org · support@instats.org

Cite the pack. Zyphur, M. J. (2026). *The Institutional AI Readiness Pack: Self-Assessment and Implementation Tools for Responsible AI in Academic Research*. Instats Policy Series. https://doi.org/10.61700/bv2nulyhht

Companion report. Zyphur, M. J. (2026). *Responsible AI in Academic Research: A Competency Framework for Research Training*. Instats Policy Series. https://doi.org/10.61700/t31oy23grr