

B5 — Briefing for graduate-student representative bodies

Instrument B5 • Briefing for graduate-student representative bodies and doctoral-student representatives

What graduate students are entitled to expect from their institution on AI in research, and how a representative body can make an institution prove it.

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

About this instrument

Purpose	Turn the five-dimension framework into entitlements a representative body can claim for graduate students, and into answerable requests to a named institutional office.
Who uses it	National and institutional graduate-student associations, doctoral and graduate-research student representatives, postdoctoral associations, and graduate-student members of graduate-studies and research-integrity committees.
Time required	Ten minutes to read. About two hours to turn Section 4 into a dated written request to one institution.
Related report sections	§1.3 (institutional baseline), §2.1–§2.5 (the five dimensions), §4.1–§4.3 (the grid and its scoring), §5.4 (adjudication), §5.5 (representative bodies), and Appendices C and G.
Grid cells it touches	D1-policy, D1-people, D1-process, D2-policy, D2-process, D3-systems, D4-people, D4-process, D5-policy, D5-people, D5-process.

Table 1. Summary of the core purpose, scope, and foundational specifications of the policy instrument.

The five dimensions. D1 Human-in-the-loop discipline · D2 Responsible use in practice · D3 Tooling that promotes responsible use · D4 AI-literate humans · D5 Institutional benchmarking grid. Each dimension is scored across four axes (policy, people, systems, and process) at one of four developmental levels: absent, nascent, established, and leading. A dimension is only as mature as its weakest axis, meaning that advancing along any dimension requires meeting the standard across all four axes.

Throughout this brief, I use the term graduate student for the general case, and doctoral student when formal enrolled status is the focus. That term covers anyone enrolled in a research doctorate at any stage, regardless of the local terminology, such as PhD students in the United States, research students in the United Kingdom, or higher degree by research students in Australia. I avoid the word candidate because in United States usage it is reserved specifically for students who have already passed their qualifying examinations. Supervisors are typically called advisors in the United States, so I write advisors and supervisors whenever either role is meant. I also use three other terms in their broadest sense: thesis refers to either a thesis or a dissertation, examiners refers to the faculty members appointed to assess it (known as dissertation committee members in the United States), and oral examination refers to the final defense of the thesis, commonly called the viva in the United Kingdom and Australia and the dissertation defense in the United States.

1. Whose document this is

I designed the Instats AI Readiness Pack primarily for institutions. Its instruments ask a university to assess and score itself across all five dimensions, name an accountable owner for each axis, and publish the results. Nothing legally obliges a university administration to complete this assessment, and that is precisely what makes it useful for student advocates. It provides representative bodies with an independent, published yardstick based on observable criteria from a credible external source. Because it was written independently of institutional bargaining, and because most of the required evidence involves public documentation, you can use it immediately.

Two factual findings from *Responsible AI in Academic Research: A Competency Framework for Research Training* establish your opening position. First, among the 38 universities across 15 countries and jurisdictions reviewed in the report's baseline study (§1.3), only six publish AI policies that explicitly address AI literacy, graduate supervision, and thesis examination. Roughly four in ten institutions stop at what I call the plagiarism ceiling, treating generative AI solely as an academic integrity issue. Second, no university currently publishes its self-scoring against this readiness framework. In other words, reaching the *leading* level on Dimension 5 is a rung on the developmental ladder that no institution currently occupies (§4.2).

Your opening position with university leadership is therefore not to claim that the institution is falling behind its peers. Instead, your position is: *no sampled institution yet publishes its own assessment against this framework, and we are asking our university to take the lead by doing so.*

2. Six things a doctoral student should be able to expect

I have framed each of these six entitlements around concrete, verifiable evidence so that a specific document either exists or does not. A university administration can easily dispute vague feelings or subjective complaints, but it cannot dispute whether a dated, written document actually exists.

#	The entitlement	The artifact that proves it	Anchor
E1	Rules in writing before the work is done. Graduate students know clearly at the outset of their research which tasks AI may assist with and which must remain strictly their own independent work.	A dated, published task-level demarcation of labor and judgment (in the form of Appendix G), rather than a single vague policy sentence stating that AI must not replace scholarly judgment.	D1 · §2.1
E2	An advisor or supervisor who has had the conversation, and a formal record of it. This should be a documented agreement rather than an informal hallway conversation that a student must try to recall four years later.	A completed, dated, and countersigned supervisor-student AI-use agreement (such as Instrument G1) held by both parties. For example, ETH Zurich already requires that AI use and source declarations be formally agreed upon with the supervisor.	D1 · §2.1, §4.2
E3	Training before judgment. Graduate students are formally taught the six AI-era research competencies before their research work is evaluated against them.	A research-methods curriculum outline that explicitly names the six core competencies as learning outcomes: citation verification, model-and-parameter specification, prompt-as-fork-in-the-garden discipline, model-heterogeneity in adversarial review, sycophancy detection and human-as-verifier discipline, and structured failure-mode reporting.	D4 · §2.4

#	The entitlement	The artifact that proves it	Anchor
E4	No retrospective reclassification. This pack protects conduct that was permitted when undertaken from reclassification solely because policy changes, subject to continuing legal, ethical, contractual, confidentiality, funder, and publisher obligations.	An institutional policy document with a clear version number, effective date, change log, and a formal transition clause stating this bounded protection.	P1 Clause 15.6 · G1 §8
E5	Proportionate handling, correctly classified. An undisclosed use of AI is not the same as data fabrication, and a fabrication is not a minor formatting error.	A written adjudication framework defining the four distinct case types from §5.4 (undisclosed use, disclosed-but-unverified use, AI-fabricated content, and AI-assisted misconduct), along with the standard of proof, range of remedies, right of reply, and formal appeal route for each.	D2 · §5.4
E6	Student data and research remain protected. No graduate student should be required, or informally encouraged, to enter unpublished research results or sensitive data into tools that train on user inputs.	An official institutional register of approved AI tools stating, for each tool, whether it is enterprise-tenanted with training-on-input disabled, where data is stored, how long session logs are retained, and who is authorized to view them.	D3 · §2.3

Table 2. Doctoral student entitlements alongside corresponding verifiable evidence artifacts and policy anchors.

3. Six risks graduate students carry that institutions routinely miss

R1: The rule that did not exist until the oral examination. The University of Toronto's School of Graduate Studies provides a clear example: AI use in researching or writing the thesis requires "the prior approval of the supervisor(s) and supervisory committee," and doctoral students must be able to "describe and defend any use of generative AI" at the final oral examination. When an institution adopts the requirement to defend AI use at examination without first establishing a formal front-end approval process, it leaves the doctoral student bearing all of the personal and academic risk with none of the necessary guidance or notice.

R2: The supervisor lottery. When central institutional policy is absent or vague, the operative rule defaults to whatever an individual advisor or supervisor happens to believe. Where policy authority is decentralized, guidance at the school or departmental level might instruct students to assume that all AI use is prohibited unless explicitly told otherwise. As a result, two doctoral students in the same university cohort can end up working under completely contradictory rules. That is a clear failure on the people axis of Dimension 1, and the practical burden of that inconsistency falls directly on the students.

R3: Pressure to use tools nobody trained them on. In many research environments, a lab adopts a new software tool and expects graduate students to be fluent in using it, even when no formal training has been provided. If the tool produces errors or hallucinations, those mistakes are often blamed on the student. Every tool class documented in Appendix C of the report carries known, documented failure modes, ranging from fabricated literature citations in search tools to invented text in audio transcription. Expecting students to navigate these tools safely without providing institutional training is a serious failure on Dimension 4, and it imposes an unfair personal and professional cost on graduate researchers.

R4: Detection tools used as evidence. Automated AI detection software is unreliable and should be treated as rough triage at best. In a study evaluating seven commercial detectors, Perkins et al. (2024) found an average baseline accuracy of only 39.5 percent, which dropped to 17.4 percent under simple adversarial techniques, along with documented bias against non-native English speakers. For this reason, King's College London explicitly bars doctoral examiners from uploading any part of a thesis into generative AI tools and from using external AI-detection software. An unverified detector score should never form the substantive basis of a misconduct allegation against a student.

R5: The graduate student's data and intellectual property. Unpublished research findings, interview transcripts, participant data, and draft dissertation chapters represent a graduate student's scholarly contribution, and they often contain confidential or proprietary information. Consumer-tier AI tools default to using input data to train their models. Harvard University's central guidelines explicitly warn researchers not to enter unpublished research results into public AI tools. Similarly, Samsung's experience in March 2023 (where three confidential data leaks occurred within twenty days, prompting an immediate corporate ban on consumer generative AI tools) demonstrated that secure data handling requires institutional procurement rather than individual caution alone. In addition, while session-log auditability can assist institutional investigations, it also records a student's entire creative and analytical process. Student representatives should ask who has access to these logs and under what specific legal or administrative grounds they may be reviewed.

R6: Automated decisions made about the graduate student. Academic admissions, learning-outcome evaluation, student monitoring, and automated exam proctoring fall directly under the Annex III high-risk classification of the European Union AI Act. For institutions operating in or serving students from the EU market, the Article 27 obligation to conduct a Fundamental Rights Impact Assessment takes effect on August 2, 2026. In Australia, the automated decision-making transparency obligations under the Privacy Act 1988 commence on December 10, 2026, covering private universities and the Australian National University. Most state and territory public universities are governed by their respective state privacy legislation. Because the rights being evaluated under these laws are the rights of graduate students, student representatives should ask to see these risk assessments and inquire whether students were formally represented in the process.

4. The advocacy asks

I recommend submitting each request in writing to a specific, named office, along with a clear deadline for reply. When you ask for policy documents that an institution should already have in place, the request is very difficult for administrators to dismiss and straightforward for student representatives to follow up on.

Drafting note. A single covering sentence can carry all eight requests: "On behalf of graduate students at [INSTITUTION NAME], we ask [COMMITTEE] to supply the following documents, or to confirm in writing that they do not yet exist, by [DATE]." The second part of that sentence is especially important, because a written confirmation that a document does not exist provides a clear and demonstrable finding on the institution's current state of readiness.

#	What to request	Ask it of	A good answer contains	Status
A1	The published labor-versus-judgment demarcation for research tasks.	Dean of the Graduate School	A dated task-level catalog in the form of Appendix G, with a clear route for schools and colleges to adapt it.	☐ Yes ☐ Partial ☐ No
A2	Task-level rules for the four modes of AI use: search, co-author, validator, and tutor.	Research-integrity office	Four distinct rule sets, rather than a single general paragraph, plus the disclosure template that accompanies them.	☐ Yes ☐ Partial ☐ No
A3	The policy's version history and transition clause.	The policy owner named on the document	A clear version number, effective date, change log, and an explicit transition clause specifying which version governs work already in progress.	☐ Yes ☐ Partial ☐ No
A4	The adjudication framework for AI-related cases.	Graduate School or research-integrity committee	The four case types, the evidentiary standard for each, available remedies, right of reply, the formal appeal route, and whether detector output is permitted as evidence.	☐ Yes ☐ Partial ☐ No
A5	The training record for advisors, supervisors, and examiners.	Graduate School and the office responsible for administering thesis examinations	Details on what training is delivered, to whom, how often, completion rates, and whether completion is mandatory.	☐ Yes ☐ Partial ☐ No
A6	The methods curriculum showing the six competencies as named outcomes.	Associate Dean for Research Training (or equivalent)	The six research competencies identified as explicit learning outcomes, along with the schedule for regular curriculum review.	☐ Yes ☐ Partial ☐ No
A7	The approved-tools register and the procurement standard behind it.	Chief Information Officer or Procurement Office	The six core tool properties applied as a procurement gate (verifiable citation, data residency, uncertainty reporting, reproducibility, auditability, and open-source or local options), along with tool-specific details on data residency, training-on-input, and log retention.	☐ Yes ☐ Partial ☐ No
A8	The Dimension 5 answer: who owns each axis, at what cadence, with which deadlines assigned.	Vice-President for Research, Deputy Vice-Chancellor (Research), or equivalent institutional leader	A named owner for policy, people, systems, and process across each dimension, a defined review cadence, and a dated plan for meeting applicable regulatory obligations.	☐ Yes ☐ Partial ☐ No

Table 3. Advocacy requests, designated recipient offices, required components of acceptable responses, and completion status.

Four answers that sound like compliance and are not

- *"We have an AI policy."* Having a single policy document only addresses one aspect of readiness. A policy document represents an entry in the grid, not the entire institutional system (§2.5). Ask

leadership which specific actions across the people, systems, and process axes that policy actually establishes.

- *"It is covered by our academic-integrity rules."* This response reflects what I describe as the plagiarism ceiling, where roughly four in ten institutions in the report's benchmark sample currently remain (§1.3). Academic integrity policies designed to prevent cheating do not address the broader research-integrity, data-governance, and competency requirements of responsible AI use.
- *"Advisors and supervisors handle it case by case."* Treating policy as purely discretionary is simply an *absent* rating on the Dimension 1 people axis, framed as administrative flexibility. It guarantees the supervisor lottery described in Risk 2.
- *"We are waiting for sector guidance."* Publishers converged on the no-AI-co-author rule within ten weeks and later developed disclosure conventions. Funders moved more slowly and unevenly (§1.1). Comprehensive frameworks like this one are already published and available. When you hear this response, ask for a specific target date by which the institution will complete its guidance.

5. How the supervisor-student agreement protects the student

Instrument G1 is often presented primarily as an institutional risk control. From the perspective of graduate students, however, it accomplishes five vital protections:

1. **It dates the rule.** A signed and dated agreement provides the student with clear, contemporaneous evidence of what was formally permitted when the research was conducted. This serves as the primary practical defense against Risk 1 (rules appearing unexpectedly at the final defense).
2. **It converts a preference into a position.** An advisor's or supervisor's unwritten, personal preference cannot be formally appealed. In contrast, a written agreement can be checked against institutional policy and corrected if it exceeds the supervisor's authority.
3. **It makes the advisor or supervisor an accountable party.** When an advisor or supervisor directs or approves a specific use of AI, the written document creates a clear record of that guidance. In other words, accountability and responsibility follow the person who made the decision.
4. **It operates at the level of specific research tasks.** Drawing directly from the task demarcation in Appendix G, the agreement explicitly distinguishes between technical labor and scholarly judgment, rather than relying on vague instructions to use good sense.
5. **It changes prospectively.** Any revision to the agreement requires an updated date and change log. Changes do not retrospectively reclassify conduct permitted when undertaken, subject to continuing legal, ethical, contractual, confidentiality, funder, and publisher obligations.

The test before endorsing an institutional agreement. Before supporting any supervisor-student agreement template at your university, ask a simple question: does the document bind both parties, or does it place obligations solely on the graduate student? A fair, balanced agreement clearly defines what the institution and the supervisor commit to providing, including access to approved AI tools, appropriate training, and a clear procedure for updating the agreement as research evolves. An agreement that records only the student's obligations while shielding the institution is simply a liability transfer with a signature block, and student representatives should feel confident calling it exactly that.

6. If you benchmark, benchmark properly

In §5.5 of the report, I encourage graduate student representative bodies, including national and system-level associations, to commission and publish independent institutional benchmarking. If your organization decides to benchmark universities, following four straightforward rules will ensure your findings remain robust and defensible:

- **Never reduce an institution to a single overall score.** The genuine diagnostic value of this framework lies in the pattern across all five dimensions rather than an uninformative composite score (§4.3). When you publish results, report the developmental level for each of the five dimensions and identify the specific axis (policy, people, systems, or process) that is holding back maturity in each area.
- **Evaluate dimensions using the lowest-scoring axis rather than an average.** To achieve an *established* level on any dimension, an institution must have the policy, people, systems, and process axes all functioning at that level (§4.1). Calculating a numerical average conceals institutional gaps and obscures the actual vulnerabilities that graduate students encounter in practice.
- **Score only what is publicly documented and verifiable.** If an institution claims to follow a practice but cannot provide a public document or verifiable policy, record that indicator as unevidenced. Always state the proportion of indicators for which you were able to locate evidence, and if your total evidence coverage is below 80 percent, label the assessment as provisional.
- **Provide a formal right of reply with a clear deadline, and publish the response.** Before publishing your assessment, share the draft findings with the institution and give administrators a reasonable window to provide any missing documentation. If an institution supplies a verified document that was previously missing, it improves both the university's score and the accuracy of your report, which is the exact developmental outcome this process is meant to encourage.

Crosswalk: item to report section or pack source

Item	Report section or pack source	Grid cell(s)
§1: Whose document this is	§5.5, §1.3, §4.2	D5-policy
E1: Rules in writing beforehand	§2.1, Appendix G	D1-policy
E2: Recorded supervisor conversation	§2.1, §4.2	D1-process, D1-people
E3: Training before judgment	§2.4, §5.3	D4-process
E4: No retrospective reclassification	Pack safeguard: P1 Clause 15.6 and G1 §8	D1-policy, D5-policy
E5: Proportionate handling	§5.4	D2-process
E6: Student data and work protection	§2.3	D3-systems
R1: The rule that appears at the oral examination	§2.1, §3.2, §5.3	D1-process
R2: The supervisor lottery	§1.3, §2.1	D1-people
R3: Untrained tool use	§2.4, Appendix C	D4-people
R4: Detection tools as evidence	Appendix C (Class 11), §5.4	D2-process

Item	Report section or pack source	Grid cell(s)
R5: Graduate student data and IP	§2.3, Appendix C (Classes 2 and 10)	D3-systems
R6: Automated decisions about graduate students	§2.3, §2.5	D5-policy, D3-systems
A1: Labor and judgment demarcation	§2.1, Appendix G	D1-policy
A2: Rules for the four AI-use modes	§2.2	D2-policy
A3: Version history and transition clause	§2.5	D5-policy
A4: Adjudication framework	§5.4	D2-process
A5: Supervisor and examiner training	§5.3, §2.4	D4-people
A6: Six competencies in the curriculum	§2.4	D4-process
A7: Approved-tools register	§2.3	D3-systems
A8: Axis owners, cadence, and deadlines	§2.5, §4.2	D5-people, D5-process
§5: The agreement from the student's side	§2.1, §2.2, Appendix G	D1-process, D2-process
§6: Benchmarking rules	§4.1, §4.3, §5.5	D5-process

Table 4. Crosswalk mapping individual policy instrument items to report sections or pack sources and corresponding grid cells.

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>

License. The pack and its instruments are licensed under [Creative Commons Attribution 4.0 International \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/). You may adapt them for institutional use with attribution.