instats

A1 — Institutional AI-Readiness Maturity Self-Assessment

This is the primary instrument of the Instats AI Readiness Pack accompanying Responsible AI in Academic Research: A Competency Framework for Research Training. *It uses sixty-eight evidence-anchored indicators to place an institution on the twenty-cell maturity grid, dimension by dimension, without reducing the institution to a single uninformative score.*

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

About this instrument

Purpose	To establish, from documentary evidence rather than opinion, where [INSTITUTION NAME] sits across each of the five framework dimensions and each of the four institutional axes, and to name the specific axis holding each dimension back.
Who completes it	A cross-functional group convened by the Vice President for Research, the Deputy Vice-Chancellor (Research), or equivalent, rather than a single office. See §1.2.
Time required	Three to four weeks of evidence assembly before the session, one half-day facilitated session, and a two-week window to resolve any items parked as unverified.
Related report sections	Part 2 (§2.1–§2.5), Part 4 (§4.1–§4.3), Appendix A (the full grid), and §1.3 (the Class A–D taxonomy, used in §6 of this instrument).
Cells touched	All twenty cells. This is the only instrument in the pack that evaluates the complete grid.
Companion instruments	Instrument A4 (the gap-analysis worksheet and sixty-rung action ladder, which lists ordered improvements where each step is a rung) consumes this instrument's output. Instruments A2 and A3 (the doctoral student and supervisor surveys) test whether the institution's paper position matches lived practice. Instrument A5 (the board one-pager) reports the result upward to leadership.

Table 1. Summary of key operational specifications and structural parameters of the assessment instrument.

---

1. About this instrument

1.1 What it is for, and why it exists

The report's Part 4 maturity grid crosses five framework dimensions with four maturity levels, and evaluates each intersection against four institutional axes: policy, people, systems, and process. This instrument turns that grid into practical questions an institution can answer about itself using documents on the table.

Regarding Dimension 5 (the institution's ability to know where it stands and act on it), the report's §4.2 describes that *"no university in the sample publishes its own scoring against this framework as of*

mid-2026, because the framework has just been published in this report, so leading on Dimension 5 is, for now, a forward-looking standard that the framework's adopters can elect to populate."

In other words, the report names a level of achievement, or "rung" on the developmental ladder, that nobody currently occupies. This instrument is how an institution reaches it. Completing A1 honestly, publishing the result with the method clearly stated, and reviewing it at a regular, declared cadence represents the substantive content of Dimension 5.

A second reason to use this tool is narrower and more immediate. Two regulatory obligations sit inside Dimension 5 with fixed compliance dates. The first is the EU AI Act Article 27 Fundamental Rights Impact Assessment obligation. This applies from 2 August 2026 to institutions within Article 2's territorial scope that are bodies governed by public law or private entities providing public services and deploy a high-risk system under Article 6(2), such as where AI touches admissions, learning-outcome evaluation, education-level assignment within the institution, or proctoring. The second is the Australian Privacy Act 1988 automated-decision-making transparency obligation, which commences on 10 December 2026 for institutions under that regime. Institutions in other jurisdictions have their own equivalents. This instrument asks whether each obligation has a named owner and a dated plan, because the report's §2.5 sets exactly that as Dimension 5's operational test: a named owner for each pressure point, a documented preparation plan, and a published timeline.

1.2 Who completes it

The assessment fails when a single office attempts to complete it alone, because no single office holds all the necessary evidence. Several cells sit outside the research portfolio altogether, in procurement, information security, and privacy. Convene the following roles, or the local equivalent of each:

- **Sponsor.** Deputy Vice-Chancellor (Research), Pro-Vice-Chancellor (Research), Vice President for Research, or Deputy Provost. This person chairs the session or formally delegates the chair.
- **Dean of the Graduate School** or Associate Dean (Research Training). Holds enrollment, milestone, and examination evidence.
- **Director of Research Integrity.** Holds the policy text, the formal adjudication record, and the research-integrity code the institution anchors on.
- **An Associate Dean (Research) from a school or college.** Represents what actually happens in a specific discipline, and can provide a realistic check if the central administrative account is overly optimistic. This perspective is very useful.
- **Doctoral supervision lead** or chair of the supervisor-development program (advisor development, in US usage). Holds training content and attendance records.
- **Academic-integrity adjudicator** or equivalent. Holds the case record and knows how AI matters are currently classified.
- **Librarian or research-information specialist.** Holds the AI-literacy provision, citation-verification teaching materials, and reference-management integrations.
- **Chief Information Security Officer or delegate.** Holds tenancy configuration records, session-log retention schedules, and the software tool register.
- **Data Protection Officer or Privacy Officer.** Holds impact assessments, records of processing, and the regulatory deadline compliance posture.
- **Procurement lead for software and AI services.** Holds the procurement gate documentation, vendor contracts, and the exception register.
- **A doctoral student representative**, nominated by the graduate student body. This representative should be present for the whole session, not only for the parts that directly concern students.

- **A facilitator** who is not one of the above, who manages the scoring process and has no personal stake in the result.

> ***Drafting note (group size).*** *A group of twelve people is workable for a half-day session, but fifteen is too large. Where roles overlap at [INSTITUTION NAME], you should combine them. Where a role does not exist, record that as a finding, because several cells in Dimension 3 and Dimension 5 test whether the role exists at all.*

1.3 Declare the unit of assessment before you start

Score the central institution unless you decide otherwise and formally record the decision. The report's worked examples show why this matters: an institution can carry substantial activity at the school and college level while its central position is weak, and averaging the two produces an uninformative number that describes nobody. Where a school or college holds instruments the center does not, log them in the notes column against the relevant cell and score the center on its own evidence.

A school or college may also run this instrument separately against itself. The report's §5.2 treats the difference between a school or college profile and the institutional profile as the most informative diagnostic available. For instance, a school or college at *established* where the institution is at *nascent* is doing work the central institution should adopt, while a school or college at *absent* where the institution is at *established* clearly needs targeted intervention.

1.4 Running it as a facilitated session

Before the session (three to four weeks). The facilitator circulates this instrument in full and asks each participant to assemble the documentary artifacts for the cells they oversee. This means assembling evidence rather than finalizing answers ahead of time. Participants should send a list of documents rather than a completed self-assessment. The point of this preparation is to make sure every relevant document is in the room on the day.

Session agenda (four hours).

Time	Item
0:00–0:15	The evidence rule and the Partial rule (§2). The facilitator explains both aloud and takes questions.
0:15–0:25	Declare the unit of assessment (§1.3), and record the decision.
0:25–1:10	Dimension 1 (fourteen indicators across four axes).
1:10–1:55	Dimension 2 (fourteen indicators).
1:55–2:10	Break.
2:10–2:50	Dimension 3 (fourteen indicators). Expect the procurement and privacy participants to lead this discussion.
2:50–3:15	Dimension 4 (thirteen indicators).
3:15–3:35	Dimension 5 (thirteen indicators).
3:35–3:50	Class A–D placement (§6), consisting of eleven questions answered from the policy text alone.
3:50–4:00	Read back the parked items, and assign an owner and a resolution date to each.

Table 2. Schedule of timed activities and discussion topics for the facilitated evaluation session.

Facilitator's rules. Only one person answers each indicator: specifically, the person who holds the direct documentary artifact. Nobody else votes. When an artifact is not physically or digitally available during the session, the answer must be recorded as *parked*, never as a *Yes*. If two participants disagree about what a document actually says, open the document and check the text together. In my experience, a rigorous session should produce a written explanatory note against roughly one indicator in five. If a session produces no notes at all, it is very likely that the group has not been sufficiently thorough.

After the session (two weeks). During the two weeks following the session, the facilitator resolves any parked items against the artifacts, records the final answers, computes the profile according to §3, and returns the completed scoring sheet with the full evidence register attached. Any parked item that remains unresolved stays unanswered, and the final coverage figure reflects that. Do not treat an inability to locate evidence as proof that the underlying practice is absent. The completed profile is then formally tabled at [COMMITTEE], which sets the date for the next assessment. I strongly recommend fixing that date before the group disperses, because a profile without a scheduled review date quickly turns into an administrative exercise that nobody actively owns.

---

2. How to answer honestly

2.1 The evidence rule

Every indicator asks about something an outsider could confirm or refute from documents. That property is what makes the result worth publishing.

The test to apply to every answer is straightforward: *could a skeptical auditor who does not work here confirm or refute this from the artifacts we hold?* If the honest answer is that they would have to take our word for it, then the indicator is not satisfied.

What counts as evidence. Examples of valid evidence include a published policy document with a specific clause reference and date, a live system screen or form, a training completion record extracted directly from an institutional database, an official committee minute, a signed procedure, a dated register entry, a formal report with a named recipient, an executed contract clause, or a completed assessment or impact-assessment artifact with an explicit completion date.

What does not count as evidence. The assessment cannot accept an unfulfilled intention, an unapproved draft, a pilot project with no completion records, an individual's informal recollection of a past decision, a slide from an old presentation, an undated intranet page, or a vendor's marketing claim about its own software. The report's §2.3 is very specific on this point: the Stanford RegLab 2024 evaluation measured seventeen to thirty-three percent hallucination rates across three commercial retrieval-augmented tools that were marketed as completely eliminating false citations. A vendor's marketing statement is a claim that needs to be tested, not an artifact that satisfies an institutional indicator.

Reference every answer to a named artifact. The evidence line under each indicator specifies the type of artifact expected, and you should fill this in completely. If an indicator is marked as a Yes but leaves the evidence line blank, it is considered unanswered, and the facilitator will treat it as such.

Accept documented functional equivalents. An institution does not need to use the document title, system name, or organizational form given in an example. A local artifact qualifies when it performs the same function, covers the same population and scope, and supplies the evidence required by the indicator. Record the equivalence in the evidence line so another reviewer can follow the judgment.

2.2 The Partial trap

Partial is the answer most often misused, because it feels safe. It is not a hedge.

You should use *Partial* only where a practice or mechanism genuinely exists but does not yet meet the full standard required by the indicator. For example, use it when a policy document is published but covers only two of the four AI-use modes, when advisor and supervisor training exists but only reaches staff in a single school or college, when an asset register exists but has not been updated in eighteen months, or when an impact assessment is complete for two of the four regulated use cases. In this framework, the gap analysis arranges needed improvements in sequence using a *ladder* metaphor, and each individual improvement is a *rung*. A Partial answer identifies unfinished work. Instrument A4 then helps the institution select a relevant planning package and name the exact A1 indicator it intends to rescore. The catalogue is not a one-to-one conversion of every Partial answer into a prewritten action.

A Partial answer never promotes an institution's level within a cell. A given rung is met only when every indicator on that rung is answered with a full Yes. If a rung contains a Partial answer and no No or unanswered items, it is recorded as *in progress* on the scoring sheet. This distinction is consequential for institutional planning, but it does not raise the cell's maturity level. This reflects what the report calls the cumulative ratchet rule, meaning that advancement builds strictly on solid foundations: every description for the *leading* level in Appendix A begins with "All of *established*, plus". As a result, an institution cannot be classified as *leading* while failing to satisfy the *established* criteria, and by the same logic, it cannot be classified as *established* while only half-meeting those requirements.

Do not use *Partial* to mean "we intend to do this". An intention is a "No". Similarly, do not use it to mean "this happens informally in some schools or colleges but is not central policy" unless the declared unit of assessment is that specific school or college. At the central institutional level, that situation is a "No", accompanied by an explanatory note naming the specific schools or colleges concerned.

2.3 Why guessing high (optimistically) defeats the purpose

There are three main reasons why guessing high (optimistically) defeats the entire purpose of this assessment, listed here in ascending order of importance.

First, the output of this instrument directly informs a practical work plan. Every No and Partial answer identifies work that must be considered. Instrument A4 provides sixty planning packages, but the institution must select and adapt the relevant package, name the exact A1 indicator being addressed, and rescore that indicator after the work. If an institution inflates an answer, necessary work may not be recognized or funded.

Second, the resulting institutional profile is designed to diagnose constraints rather than serve as a generic report card. The report's §4.3 makes this point plainly: the specific dimension where an institution is weakest is often the bottleneck that constrains its overall competency. For example, an institution might appear to be *leading* on Dimension 1 (human-in-the-loop policy), but if it remains *absent* on Dimension 3 (procured systems and tooling), it cannot realize the human-in-the-loop discipline in practice because the software tools its researchers actually use do not support it. Inflating the score on Dimension 3 hides the practical constraint that is already undermining Dimension 1. The grid exists to make these institutional bottlenecks visible, and an artificially flattered grid is worse than having no diagnostic at all.

Third, published scores are publicly verifiable. Moving to the *leading* level on Dimension 5 requires an institution to publish its self-assessment along with its methodology. This allows peer universities, government regulators, research funders, or doctoral students to evaluate the claim against the same objective criteria. An institution that publishes an inflated assessment creates an

easily auditable falsehood. It is far safer, and much more useful developmentally, to publish a *nascent* rating accompanied by a dated action plan than to claim an *established* rating without the necessary documentary evidence.

A related point to keep in mind is that nobody outside the institution is grading this assessment. The report's §4.1 is explicitly non-prescriptive about which maturity level any given institution ought to occupy, because institutional missions and contexts vary widely. For instance, the report notes that a small, humanities-focused doctoral program and a major biomedical research university operating across thirteen regulatory jurisdictions can both be at the *established* level on Dimension 2, even though their operational implementations look very different. There is no prize for claiming a high score and no penalty for reporting a low one, but there is a real cost to producing an inaccurate assessment.

3. How scoring works

3.1 The answer scale

Each indicator is answered **Yes (1.0)**, **Partial (0.5)**, or **No (0.0)**. Use No when the evidence establishes that the requirement is not met. If the evidence is unavailable or unresolved, leave the item blank. An unanswered indicator contributes no points and blocks confirmation of its rung, but it does not establish that the practice is absent. The denominator for documentary coverage remains fixed at sixty-eight indicators.

There is no general Not applicable category. An applicability-conditioned indicator can be answered Yes when a documented scope determination establishes why no relevant obligation or use applies to the declared unit of assessment. Record the decision, its authority, date, scope, and rationale in the evidence line. Without that record, the indicator remains unanswered rather than being removed from the denominator.

3.2 From indicators to a cell level

Each cell carries indicators across three developmental rungs: **nascent (N)**, **established (E)**, and **leading (L)**. The *absent* level has no explicit indicators. Assign it only when the nascent rung is documented as not met or is fully answered but in progress. If an unanswered indicator makes the nascent rung unresolved, record the cell as **Unresolved/not demonstrated** and do not assign a maturity level to that cell.

- A rung is **met** when every indicator on that rung is answered Yes.
- A rung is **in progress** when at least one indicator is Yes or Partial, none is No or unanswered, and the rung is not fully met.
- A rung is **not met** when documentary evidence supports one or more No answers.
- A rung is **unresolved** when an unanswered indicator prevents a determination and no documented No already establishes that the rung is not met.

The cell level is then the highest rung at which that rung *and every rung below it* is met:

Cell level	Condition
Absent	The nascent rung is documented as not met or is in progress.
Nascent	The nascent rung is met, but the established rung is not.
Established	The nascent and established rungs are both met.
Leading	The nascent, established, and leading rungs are all met.

Table 3. Qualifying conditions and progression criteria required for achieving each cell maturity level.

The cumulative ratchet rule is built directly into this scoring structure: satisfying a leading indicator while failing an established one does not raise the cell's maturity level. You should record the leading answer in the evidence register anyway, because it documents good work and will shorten the action ladder later, but it does not move the institution's location within a given cell.

If an unanswered indicator is what prevents the nascent rung from being demonstrated, do not assign *absent* on the basis of that blank. Mark the cell **Unresolved/not demonstrated** alongside the provisional documentary profile. This is an evidence-status notation, not a fifth maturity level. If a documented No independently establishes that the nascent rung is not met, the cell may still be recorded as *absent*, with the unanswered indicator reported separately. For an unresolved higher rung, retain the highest lower level that has been fully demonstrated and mark the next rung unresolved.

3.3 From cells to a dimension level

The dimension level is the minimum of its four axis cells. Not the mean, not the median nor the mode. The report is explicit that *established* means policy, people, systems, and process are all in place and operating at a reviewable cadence. As a result, a dimension with three cells at *established* and one at *nascent* remains at *nascent* for that dimension. Advancing along a dimension requires every cell to reach that level.

If any axis cell is **Unresolved/not demonstrated** at the nascent rung, report the dimension as **Unresolved/not demonstrated** in the provisional profile and list the unresolved axes. Do not assign a binding axis or binding dimension until the missing evidence is resolved, unless a documented lower cell already determines the dimension's minimum independently. Documentary coverage is reported separately. A coverage verdict of Final therefore does not turn an unresolved maturity result into a final maturity level.

Name the binding axis. The binding axis is the specific institutional axis (policy, people, systems, or process) that is holding a dimension back, meaning it sits at the lowest level. If two or more axes tie at the minimum level, you should name all of them. Identifying the binding axis is one of the most practically useful outputs of this assessment because it turns a score into a clear operational instruction. For example, if Dimension 4 sits at *nascent* with a binding axis of *people*, this tells you that the formal curriculum is not the problem. Instead, the preparation of advisors, supervisors, and examiners (or dissertation committee members, in US usage) is the problem that needs addressing.

Name the binding dimension. Across the five dimensions, the binding dimension is the one sitting at the lowest level. If two or more dimensions tie, name the earliest one in the framework sequence. The report's §4.1 explains that the five dimensions are interdependent in the order specified by the framework, operating as sequential preconditions. To see why the dimensions work this way, consider how each one depends on the one before it. An institution that has not decided where human judgement must remain mandatory cannot define what responsible use looks like in daily practice. Without a clear operational definition of responsible use, any procurement test to determine whether software promotes responsible behavior is meaningless. Without approved tools and operational standards, an AI-literacy curriculum has no concrete practices to teach. And without all four of those components in place, an institutional benchmarking grid has nothing solid to evaluate.

3.4 The instrument refuses to produce a single institutional score

This instrument does not compute an overall score, a percentage, a star rating, or a grade, and no adaptation of it should. This design choice is not a matter of modesty. It is a direct conclusion from the report's research findings.

The report's §4.3 concludes from its worked examples that an institution's maturity is rarely uniform across dimensions, and that *the diagnostic value of the grid is the dimensional pattern, not a single score*. For example, the report places University College London at *established* moving to *leading* on Dimension 1, but at *nascent* on Dimension 3. Averaging those two ratings into a single number would erase both realities and describe an institution that does not actually exist. An averaged score would also hide the specific bottleneck holding the institution back, and identifying that bottleneck is the entire practical point of running this assessment.

A single aggregate score also encourages misleading comparisons between institutions with completely different research profiles, disciplinary mixes, and regulatory obligations, using a scale that was never intended for that purpose. Furthermore, it tempts institutional leadership to optimize a single summary number rather than fixing the specific weak cells that create risk. For that reason, this assessment reports the five distinct dimension levels, their binding axes, and the binding dimension. If senior leadership requests a single summary metric, provide them with the five dimension levels and the binding dimension, which fit easily on a single line.

3.5 Coverage and the status of the verdict

Coverage is the proportion of the sixty-eight indicators answered Yes, Partial, or No from a named documentary artifact. Parked items and blank entries are treated as unanswered.

- **Coverage of 80 percent or more:** The assessment verdict is **Final**. Record the exact coverage percentage alongside the verdict.
- **Coverage below 80 percent:** The assessment verdict is **Provisional**. Report the resulting levels, mark them as Provisional, name the specific cells carrying the unanswered indicators, and set a firm deadline to complete them.

A Provisional verdict is a completely legitimate first-year result. It is not a failure. Rather, it provides an accurate measure of how accessible the institution's own documentary evidence is, which is itself an important systems finding for Dimension 5.

---

4. The instrument

When completing this instrument, answer each indicator directly from documentary artifacts. If the relevant artifact is not available during the session, record the item as parked rather than guessing.

Dimension 1: Human-in-the-loop discipline

This dimension reflects the institutional commitment that the judgment steps defining research remain human. Labor-intensive steps may be augmented by AI, but core judgment steps cannot be outsourced to AI without ceasing to be genuine research. See the report's §2.1, Appendix A row 1, and Appendix G, which supplies a working task taxonomy.

D1 · policy

D1-policy-N1 · nascent. Has [INSTITUTION NAME] published a position (in a document that sits outside its student-conduct and assessment-integrity rules) stating that AI-generated content must be verified before use and does not replace expert opinion or judgment?

> ☐ Yes ☐ Partial ☐ No

Evidence: Name the document and the specific clause, provide the URL, state the date of last review, and name the policy-library category under which it is filed.

D1-policy-E1 · established. Does that document publish a task-level demarcation (a named catalog of research tasks classified as work AI may support versus work AI may not substitute for) rather than relying on a statement of principle alone?

☐ Yes ☐ Partial ☐ No

Evidence: Attach the task table or list with its clause reference, and state how many research tasks it classifies.

D1-policy-E2 · established. Does the demarcation name the anchor instrument through which it is enforced (such as a national research-integrity code, an institutional sworn declaration or affidavit, or the declaration form attached to the thesis or dissertation) together with a regular review cadence, and did the most recent review take place on schedule?

☐ Yes ☐ Partial ☐ No

Evidence: The clause naming the anchor instrument, the published review schedule, and the version note or committee minute recording the last review date.

D1-policy-L1 · leading. Has the institution published its own scored position on this dimension, with the scoring method clearly stated and the date of the next scheduled review published?

☐ Yes ☐ Partial ☐ No

Evidence: URL of the published scoring, the method statement, the publication date, and the next review date.

D1 · people

D1-people-N1 · nascent. Is a single named role accountable for the labor-versus-judgment demarcation, meaning a position with a specific title and a current occupant rather than a standing committee described in general terms?

☐ Yes ☐ Partial ☐ No

Evidence: Role title, current occupant, and the terms of reference or formal delegation instrument that assigns this accountability.

D1-people-E1 · established. Does the development program for doctoral advisors and supervisors include the demarcation as a taught element, with attendance formally recorded against the register of academics currently advising or supervising doctoral students?

☐ Yes ☐ Partial ☐ No

Evidence: The training materials or session outline, the attendance extract, and the proportion of currently active advisors and supervisors who completed the training in the most recent cycle.

D1-people-L1 · leading. Does the institution publish coverage figures for this dimension (specifically the proportion of advisors, supervisors, and examiners prepared, along with the required refresh interval), and can it name a concrete change made in response to those figures?

☐ Yes ☐ Partial ☐ No

Evidence: The published coverage figures, and the committee paper or minute recording the resulting institutional change.

D1 · systems

D1-systems-N1 · nascent. Does the thesis-submission or doctoral-enrollment system contain a dedicated field in which the student records generative AI use?

> ☐ Yes ☐ Partial ☐ No

Evidence: The form or system screen, along with the date the field was introduced.

D1-systems-E1 · established. Does that system field require the student to record AI use against the published task categories, rather than accepting an unguided free-text note or a simple yes/no acknowledgment?

> ☐ Yes ☐ Partial ☐ No

Evidence: The form showing the structured task categories, and the completion instructions issued to students.

D1-systems-L1 · leading. Does the system automatically generate declaration-completion rates and examiner-flag rates as a standing report to a named committee?

> ☐ Yes ☐ Partial ☐ No

Evidence: The standing report, the receiving committee, the reporting frequency, and the most recent figures.

D1 · process

D1-process-N1 · nascent. Is the demarcation formally discussed with the doctoral student at a defined milestone in the program (such as confirmation, mid-program review, or an equivalent milestone) rather than left entirely to individual advisor or supervisor discretion?

> ☐ Yes ☐ Partial ☐ No

Evidence: The milestone form or procedure text that mandates this conversation, and the date it was introduced.

D1-process-E1 · established. Does the doctoral research methods curriculum teach the demarcation as a named learning outcome, complete with a stated assessment point?

> ☐ Yes ☐ Partial ☐ No

Evidence: The unit or module syllabus showing the learning outcome and how it is assessed, along with the student cohorts it currently reaches.

D1-process-E2 · established. Are examiners provided with the demarcation as part of their standard appointment pack, and does the examination procedure explain how to raise an AI-related question or concern without it being treated as a misconduct allegation by default?

> ☐ Yes ☐ Partial ☐ No

Evidence: The examiner appointment pack showing the date the AI section was added, the distribution record for the most recent examination round, and the escalation clause in the examination regulations.

D1-process-L1 · leading. Does the institution run a periodic audit of doctoral advising and supervision that samples individual student enrollments for evidence that the demarcation was applied in practice, and are those findings published?

> ☐ Yes ☐ Partial ☐ No

Evidence: The audit protocol, the sample size, the most recent findings, and where they were published.

Dimension 2: Responsible use in practice

This dimension evaluates how an institution operationalizes "responsible AI use" from general principles into daily research workflows, specified separately for each of the four AI-use modes: search, co-author, validator, and tutor. See the report's §2.2 and Appendix A row 2.

D2 · policy

D2-policy-N1 · nascent. Has the institution published a single standard disclosure template, or an approved form of words, for recording AI use in research outputs across all schools and colleges?

> ☐ Yes ☐ Partial ☐ No

Evidence: The template, its URL, and the policy clause that mandates its use.

D2-policy-E1 · established. Do the published rules specify what responsible use means separately for each of the four AI-use modes (search, co-author, validator, and tutor) at the level of specific research tasks rather than as high-level generalities?

> ☐ Yes ☐ Partial ☐ No

Evidence: The rules document, with the clause covering each of the four modes identified individually.

D2-policy-E2 · established. Does the disclosure standard explain how the institutional template maps onto the various publisher conventions researchers encounter upon submission, such as a declaration section above the references, a methods-section note, an acknowledgments entry, or a structured submission-form field?

> ☐ Yes ☐ Partial ☐ No

Evidence: The mapping table or guidance note, and the specific publisher policies it references.

D2-policy-L1 · leading. Does the institution review its four-mode rules at a declared cadence against the evolving landscape of publisher and funder policies, and does it publish what changed at each review?

> ☐ Yes ☐ Partial ☐ No

Evidence: The review schedule, the formal record of the last review, and the published change log.

D2 · people

D2-people-N1 · nascent. Is a named office responsible for maintaining the AI disclosure template and answering researcher inquiries about it, with a clearly published contact pathway?

> ☐ Yes ☐ Partial ☐ No

Evidence: The office name, the contact route, and the webpage that publishes it.

D2-people-E1 · established. Are the four-mode rules actively taught to doctoral students by designated instructors as part of the formal research-training program, rather than merely circulated as a passive document?

☐ Yes ☐ Partial ☐ No

Evidence: The training session or course module, the delivering unit, and the enrollment or attendance figures for the last cycle.

D2-people-L1 · leading. Does the institution publish annual evidence showing whether these rules are being followed in practice, such as disclosure-completion rates, citation-verification audit results, or examiner-flag analyses?

☐ Yes ☐ Partial ☐ No

Evidence: The published figures, the methodology used to produce them, and the publication date.

D2 · systems

D2-systems-N1 · nascent. Is the disclosure template embedded directly inside the digital systems researchers and graduate students use (such as thesis submission portals, repository deposit systems, or ethics applications) rather than only provided as a downloadable document?

☐ Yes ☐ Partial ☐ No

Evidence: The system screen showing the embedded template or data field.

D2-systems-E1 · established. Does the institutional repository or thesis archive store the AI disclosure as queryable structured data, rather than burying it as unstructured free text inside the thesis file?

☐ Yes ☐ Partial ☐ No

Evidence: The metadata field definition, and a sample query result for the most recent deposit year.

D2-systems-L1 · leading. Can the institution extract an overall disclosure-completion rate for a given year directly from its administrative systems, and is that figure formally reported to a named committee?

☐ Yes ☐ Partial ☐ No

Evidence: The formal report, the reported figure, the receiving committee, and the date it was tabled.

D2 · process

D2-process-N1 · nascent. Is disclosure of AI use required at a defined checkpoint in the research lifecycle (such as an ethics application, annual progress milestone, or thesis submission) rather than left to the researcher's personal discretion?

☐ Yes ☐ Partial ☐ No

Evidence: The form or procedure that asks the question, and its location within the research workflow.

D2-process-E1 · established. Does standard supervisory practice require the advisor or supervisor to verify the student's disclosure against the actual research work undertaken, with that verification formally recorded?

☐ Yes ☐ Partial ☐ No

Evidence: The supervision record template showing the verification checkpoint, and a sample of completed records from the current academic year.

D2-process-E2 · established. Does the examination protocol require the doctoral candidate to be able to explain and defend any AI use during the oral defense, and does the examiner report form provide a specific section to record that this took place?

☐ Yes ☐ Partial ☐ No

Evidence: The examination regulation clause, and the examiner report form showing the dedicated field.

D2-process-L1 · leading. Does the institution contribute its compliance and disclosure data to a sector-level benchmarking initiative, peer consortium, or public research dataset, rather than retaining it only internally?

☐ Yes ☐ Partial ☐ No

Evidence: The name of the benchmarking exercise or peer group, the data contribution made, and the date of submission.

Dimension 3: Tooling that promotes responsible use

This dimension reflects the institutional commitment to select, procure, and deploy AI tools that support responsible use by design rather than by good intentions alone, evaluated against six observable properties. See the report's §2.3, Appendix A row 3, and Appendix C, which describes the eleven software tool classes the standard must cover.

D3 · policy

D3-policy-N1 · nascent. Does a published institutional procurement pathway exist through which AI tools intended for research must pass before enterprise deployment?

☐ Yes ☐ Partial ☐ No

Evidence: The procurement policy clause covering AI tools, and the published entry point for the procurement gate.

D3-policy-E1 · established. Does the published procurement standard evaluate all six observable properties: verifiable citation and resolvability, data residency and governance over training-on-input, first-class uncertainty reporting, reproducibility at a known model version, auditability via session logs, and open-source or local options where the research workload supports them?

☐ Yes ☐ Partial ☐ No

Evidence: The procurement standard, identifying the specific clause covering each of the six properties. Answer Partial if fewer than six are covered, and list the missing properties.

D3-policy-E2 · established. Does the standard include a documented exception process for specialized research workloads that genuinely require consumer-grade tools, clearly naming who approves an exception and under what conditions?

☐ Yes ☐ Partial ☐ No

Evidence: The exception clause, the designated approver role, and the register of exceptions granted over the last twelve months.

D3-policy-L1 · leading. Does the institution publish its tooling-procurement audit at a regular, stated cadence, and does the published audit explicitly name the software tools sampled?

☐ Yes ☐ Partial ☐ No

Evidence: The published audit, its stated cadence, the sample list of tools, and the publication date.

D3 · people

D3-people-N1 · nascent. Is a named role held accountable for deciding whether an AI tool may be approved for research use, as distinct from general administrative software procurement?

☐ Yes ☐ Partial ☐ No

Evidence: The role title, the current occupant, and the formal delegation instrument.

D3-people-E1 · established. Does the review panel include the data protection officer, the research integrity officer, and at least one research-active academic, with the committee membership publicly listed?

☐ Yes ☐ Partial ☐ No

Evidence: The terms of reference and the current published membership list.

D3-people-L1 · leading. Does the institution commission independent testing of vendor performance claims on a sample of deployed tools, rather than relying exclusively on vendor sales documentation?

☐ Yes ☐ Partial ☐ No

Evidence: The commissioning record, the external verifier, the evaluation report, and the specific tools evaluated.

D3 · systems

D3-systems-N1 · nascent. Is there an actively maintained institutional register of AI tools approved for research use that is accessible to all researchers?

☐ Yes ☐ Partial ☐ No

Evidence: The register, its location, and the date it was last updated.

D3-systems-E1 · established. Are enterprise-tenanted deployments with training-on-input disabled available for research projects involving unpublished or confidential data, and does the tool register state the data-residency status for each tool?

☐ Yes ☐ Partial ☐ No

Evidence: The tenancy configuration record, and the register entries showing residency and training-on-input status for each tool.

D3-systems-E2 · established. Does the institution retain AI interaction session logs for a clearly articulated retention window, and is that retention schedule published?

☐ Yes ☐ Partial ☐ No

Evidence: The records-governance schedule entry, the retention period, and the list of systems it covers.

D3-systems-L1 · leading. Does the institution periodically re-test deployed tools against standard citation and source resolvability tests on a stated cadence, recording the empirical result for each tool at each specific model version?

☐ Yes ☐ Partial ☐ No

Evidence: The re-testing schedule, the most recent evaluation results, and the specific model versions tested.

D3 · process

D3-process-N1 · nascent. Does a researcher who wishes to evaluate a new AI tool on research data have a published pathway to request approval, accompanied by a stated service turnaround time?

☐ Yes ☐ Partial ☐ No

Evidence: The request form or service-catalog entry, and the published service level agreement.

D3-process-E1 · established. Is the formal impact assessment required by the institution's relevant jurisdictions complete for every regulated use? For institutions operating in or serving the EU market, this requires a Fundamental Rights Impact Assessment under EU AI Act Article 27 for high-risk applications including admissions, learning-outcome evaluation, education-level assignment within the institution, and proctoring, registered as required against the 2 August 2026 application date. For institutions under the Australian legal regime, this requires the automated-decision-making disclosure posture mandated from 10 December 2026. For other jurisdictions, name and complete the local statutory equivalent. A dated, authorized determination that no relevant statutory regime or regulated use applies satisfies this indicator under §3.1 when its authority, scope, and rationale are recorded in the evidence line.

☐ Yes ☐ Partial ☐ No

Evidence: The completed assessments, the official registration record where applicable, the completion dates, and the list of regulated use cases covered by each assessment.

D3-process-L1 · leading. Does the institution contribute its empirical tool-resolvability test results to a sector-level reference benchmark or shared evaluation consortium, rather than keeping the findings confidential internally?

☐ Yes ☐ Partial ☐ No

Evidence: The named benchmark or consortium, the data contribution made, and the date of submission.

> ***Drafting note (jurisdiction).*** *Indicator D3-process-E1 is designed to be answered once, against whichever regulatory regime binds [INSTITUTION NAME]. An institution operating across multiple jurisdictions can answer Yes only when every applicable regime is fully satisfied, and it should name each regime in the evidence line. If no relevant statutory regime or regulated use applies, record the dated, authorized scope determination described in §3.1. An institution may also adopt a voluntary standard, but doing so is not required to satisfy this indicator.*

Dimension 4: AI-literate humans

This dimension evaluates the institutional commitment to ensure that doctoral researchers, their advisors and supervisors, their examiners, and research-training staff possess the practical competencies required for AI-augmented research. See the report's §2.4 and Appendix A row 4.

D4 · policy

D4-policy-N1 · nascent. Does a published institutional document state that doctoral students and their advisors or supervisors are expected to maintain specific AI-related research competencies, explicitly naming those competencies rather than referring to AI literacy in vague terms?

☐ Yes ☐ Partial ☐ No

Evidence: The policy document, the relevant clause, and the list of competencies as published.

D4-policy-E1 · established. Are the six core competencies (citation verification, model and parameter specification, prompt design as an analytical choice, model heterogeneity in adversarial review, sycophancy detection with human verification, and structured failure-mode reporting) established as required learning outcomes in the doctoral research methods curriculum, framed as essential additions to the existing canon of research validity and reproducibility rather than replacements for it?

☐ Yes ☐ Partial ☐ No

Evidence: The curriculum syllabus showing each of the six learning outcomes, along with the text establishing their relationship to traditional research methods. Answer Partial if fewer than six are named, and list the missing ones.

D4-policy-L1 · leading. Does the institution publish evidence on provision, reach, and selected aligned performance outcomes from the approved competency curriculum, on a regular, published schedule?

☐ Yes ☐ Partial ☐ No

Evidence: The published provision and participation figures, the aligned assessment method and selected performance results, and the publication date. Graduate surveys, thesis disclosures, and advisor self-assessments may describe reach, experience, or practice, but they do not by themselves demonstrate competency acquisition.

D4 · people

D4-people-N1 · nascent. Is there an institutional AI-literacy training offering available to doctoral students (such as an online course, a library workshop series, or regular seminars) with a clearly named operational owner?

☐ Yes ☐ Partial ☐ No

Evidence: The training program, its designated owner, and student participation figures from the most recent cycle.

D4-people-E1 · established. Is that AI training a mandatory component of the doctoral curriculum rather than an optional elective, with formal completion recorded on each student's official enrollment record?

☐ Yes ☐ Partial ☐ No

Evidence: The academic program regulation establishing the requirement, along with completion rates for the most recent cohort.

D4-people-E2 · established. Do advisors and supervisors receive professional development covering the same six core competencies, and do dissertation examiners receive targeted training on the corresponding verification skills?

☐ Yes ☐ Partial ☐ No

Evidence: Both training programs, a curriculum map comparing them against the six competencies, and separate completion figures for advisors and examiners.

D4-people-L1 · leading. Does the institution run an ongoing curriculum-improvement cycle for this training informed by external peer review, and can it point to a concrete curricular change resulting from that review?

☐ Yes ☐ Partial ☐ No

Evidence: The external review report, the continuous improvement documentation, and the specific curricular changes made.

D4 · systems

D4-systems-N1 · nascent. Is AI-literacy instruction delivered through a system that officially tracks completion for each individual student, advisor, or supervisor, rather than offered on an untracked open-access website?

☐ Yes ☐ Partial ☐ No

Evidence: The learning management system or administrative training record, and an extract showing verified completions.

D4-systems-E1 · established. Do institutional systems support delivery of the approved competency curriculum through versioned content, aligned assessments, and retrievable completion records for the relevant student and staff populations?

☐ Yes ☐ Partial ☐ No

Evidence: The supported delivery platform, approved curriculum version, assessment mapping, and retrievable completion records. A corpus-grounded tutoring environment may support delivery, but it is optional and is not required to satisfy this indicator.

D4-systems-L1 · leading. Can the institution generate automated reports on competency coverage by student cohort, by academic department, and across the advisor population directly from its administrative databases, without needing a manual survey?

☐ Yes ☐ Partial ☐ No

Evidence: The automated report definition and its most recent data output.

D4 · process

D4-process-N1 · nascent. Is AI literacy formally introduced during orientation for new doctoral students and during onboarding for new advisors and supervisors?

> ☐ Yes ☐ Partial ☐ No

Evidence: The student orientation agenda and the advisor appointment onboarding checklist, both dated.

D4-process-E1 · established. Is the AI methods curriculum reviewed at a regular, stated cadence to keep pace with developments in AI tooling, with the reviewing committee named and the last review meeting dated?

> ☐ Yes ☐ Partial ☐ No

Evidence: The review schedule, the minutes from the most recent review meeting, and the resulting curricular updates.

D4-process-L1 · leading. Does the institution share its competency-outcome data with sector-level benchmarking initiatives rather than keeping the information purely internal?

> ☐ Yes ☐ Partial ☐ No

Evidence: The benchmarking initiative named, the specific data submitted, and the date of contribution.

Dimension 5: Institutional benchmarking grid

This dimension measures the institution's capacity to evaluate its maturity across the four preceding dimensions and take strategic action on the findings, treating all four axes as an integrated system. See the report's §2.5 and Appendix A row 5.

D5 · policy

D5-policy-N1 · nascent. Does a published institutional document formally commit [INSTITUTION NAME] to evaluating its own maturity regarding AI in research and research training, naming the specific governance body that oversees the assessment?

> ☐ Yes ☐ Partial ☐ No

Evidence: The governing document, the relevant clause, and the designated oversight body.

D5-policy-E1 · established. Does the institution formally evaluate itself against a recognized AI maturity framework across all four preceding dimensions at a stated cadence, with the most recent evaluation dated?

> ☐ Yes ☐ Partial ☐ No

Evidence: The completed evaluation document, its completion date, and the published evaluation schedule.

D5-policy-E2 · established. Is there a documented preparation plan for each regulatory obligation applicable to the institution's jurisdictions, complete with a named owner and milestone dates for each?

> ☐ Yes ☐ Partial ☐ No

Evidence: The plan, the assigned owner for each obligation, and the specific milestone deadlines. Name each regulatory obligation that applies to the institution. For instance, the report highlights two: the EU AI Act Article 27 Fundamental Rights Impact Assessment obligation, applying from 2 August 2026 to institutions operating in or serving the EU market, and the Australian Privacy Act 1988 automated-decision-making

transparency obligation, commencing 10 December 2026 for institutions subject to that regime. Other jurisdictions have their own statutory deadlines.

D5-policy-L1 · leading. Does the institution publish its self-assessment scores along with an explicit description of its methodology, allowing an independent third party to reproduce and verify the evaluation?

> ☐ Yes ☐ Partial ☐ No

Evidence: The published assessment, the methodology statement, the public URL, and the publication date.

D5 · people

D5-people-N1 · nascent. Is there an assigned individual owner for the policy axis, meaning a specific person with a job title who is held accountable for the AI-in-research policy document and its periodic review?

> ☐ Yes ☐ Partial ☐ No

Evidence: The job title, the current role occupant, and the formal instrument assigning this accountability.

D5-people-E1 · established. Is a named operational owner assigned to each of the four axes across all five dimensions (covering all twenty cells, which may be distributed among a smaller group of individuals), and is that ownership matrix published?

> ☐ Yes ☐ Partial ☐ No

Evidence: The responsibility matrix covering all twenty cells, along with the date it was last verified.

D5-people-L1 · leading. Does the institution submit its maturity scoring to external peer review or independent audit at a stated cadence, with the evaluating organization publicly identified?

> ☐ Yes ☐ Partial ☐ No

Evidence: The external review agreement, the evaluating body, the date of the last review, and the resulting report.

D5 · systems

D5-systems-N1 · nascent. Is the documentary evidence supporting the institution's AI readiness maintained in a single central repository, rather than gathered piecemeal from different departments upon request?

> ☐ Yes ☐ Partial ☐ No

Evidence: The central register or repository, along with its index of contents.

D5-systems-E1 · established. Can the institution quickly retrieve, from a single system, the core artifacts an auditor or regulator would request, including impact assessments, the approved AI tool register, session-log retention policies, disclosure records, and training completion files?

> ☐ Yes ☐ Partial ☐ No

Evidence: The central register, and the documented results of a retrieval audit recording the specific artifacts produced and the time taken.

D5-systems-L1 · leading. Does the institution publish its maturity scoring as a machine-readable, versioned data extract that identifies each cell, its level, its evidence references, and the assessment date?

☐ Yes ☐ Partial ☐ No

Evidence: The published data extract, its schema documentation, and its version history.

D5 · process

D5-process-N1 · nascent. Has the institution formally documented its applicable AI regulatory obligations and deadlines in a paper presented to a central governing body, even if operational owners have not yet been assigned to each obligation?

☐ Yes ☐ Partial ☐ No

Evidence: The committee paper, the receiving governance body, and the meeting date.

D5-process-E1 · established. Are the self-assessment results, together with the regulatory compliance plan, formally tabled at the university council (or governing board), at school and college research committees, and at the research integrity committee on a declared schedule?

☐ Yes ☐ Partial ☐ No

Evidence: The committee papers and meeting dates for each of the three governance bodies separately.

D5-process-L1 · leading. Does the institution contribute its assessment methods and evaluation findings to sector-level standardization efforts, such as higher education representative bodies, national policy consultations, or published methodological white papers?

☐ Yes ☐ Partial ☐ No

Evidence: The specific contribution, the recipient organization, and the date of submission.

---

5. Scoring sheet

5.1 Transferring answers to the workbook

The companion workbook carries one row per indicator on the entry tab, in the order printed above. Enter Yes, Partial, or No in the answer column and the artifact reference in the evidence column. The workbook automatically computes the rung results, the twenty cell levels, the five dimension levels, the binding axes, the binary Dimension 5 assurance status, and the coverage figure. Nothing else needs to be entered.

Where the assessment is completed on paper, transcribe your answers into the two summary tables below.

Rule reminders when transcribing. A developmental rung is met only when every indicator on that rung is answered with a full Yes. A Partial answer never promotes an institution's level within a cell. An unanswered indicator contributes no points, remains in the fixed denominator for documentary coverage, and blocks confirmation of its rung, but it is not a documented No. Any leading-level answers recorded above an unmet established rung should be documented in the evidence register, but they do not raise the cell's maturity level until the established rung is fully satisfied.

5.2 Cell-level sheet

Cell	Nascent rung	Established rung	Leading rung	Cell level	Binding evidence gap
D1-policy					
D1-people					
D1-systems					
D1-process					
D2-policy					
D2-people					
D2-systems					
D2-process					
D3-policy					
D3-people					
D3-systems					
D3-process					
D4-policy					
D4-people					
D4-systems					
D4-process					
D5-policy					
D5-people					
D5-systems					
D5-process					

Table 4. Worksheet for recording rung evaluations, cell levels, and binding evidence gaps.

Rung columns take one of: met · in progress · not met · unresolved. Cell level takes one of: Absent · Nascent · Established · Leading, or records the evidence status Unresolved/not demonstrated when the nascent rung cannot be determined. Unresolved/not demonstrated is not a fifth maturity level. The binding evidence gap names the single artifact whose absence is doing the most damage in that cell.

5.3 Dimension-level sheet

Dimension	policy	people	systems	process	Dimension level (minimum)	Binding axis
D1 Human-in-the-loop discipline						
D2 Responsible use in practice						

Dimension	policy	people	systems	process	Dimension level (minimum)	Binding axis
D3 Tooling that promotes responsible use						
D4 AI-literate humans						
D5 Institutional benchmarking grid						

Table 5. Summary sheet showing institutional axis scores, overall dimension levels, and binding axes.

Dimension level takes one of: Absent · Nascent · Established · Leading, or the evidence status Unresolved/not demonstrated. Use that evidence status when an unresolved nascent rung prevents one or more axis cells from receiving a maturity level. In that case, list the unresolved axes and leave the binding dimension unassigned until the evidence is resolved, unless another documented cell already fixes the minimum independently.

5.4 The profile readout

Record these seven lines, and no others, as the formal result of the assessment.

- **Assessment date:** [DATE] · **Unit of assessment:** [INSTITUTION NAME] (central) or [SCHOOL OR COLLEGE]
- **Dimension levels:** D1 ___ · D2 ___ · D3 ___ · D4 ___ · D5 ___
- **Binding axis by dimension:** D1 ___ · D2 ___ · D3 ___ · D4 ___ · D5 ___
- **Binding dimension:** ___ (the lowest level, and on a tie, the earliest dimension in the framework sequence)
- **Dimension 5 assurance status:** Assured / Unassured (record **Assured** only when Dimension 5 is *Established* or *Leading*, and otherwise record **Unassured**)
- **Coverage:** ___ percent · **Verdict:** Final / Provisional
- **Received by:** [COMMITTEE] on [DATE] · **Next assessment due:** [DATE]

The last line is not mere administrative detail. An assessment profile with no receiving committee and no scheduled follow-up date is just an academic exercise. In contrast, a profile formally tabled at a named committee with a declared review schedule represents the actual substance of Dimension 5, and indicators D5-policy-E1 and D5-process-E1 are answered directly from it.

5.5 How to read the profile

Read the binding dimension first. This dimension represents your primary institutional constraint. Spending time and resources elsewhere will yield lower returns until this constraint is addressed, because the framework's dimensions are designed as sequential preconditions for one another.

Read the binding axis second. Identifying the binding axis turns an abstract maturity level into a practical operational instruction. The four axes tend to fall short in predictable ways, and each requires a different remedy:

- **When Policy is binding:** Staff and researchers may already be doing the work informally, but they lack an approved institutional document that defines or authorizes it. The remedy here is policy drafting, which is typically the fastest and least expensive improvement on the ladder. You can draw directly on the model policy in companion instrument P1 and the disclosure standard in P2.
- **When People is binding:** A policy document exists on paper, but no specific individual is accountable for implementing or maintaining it. This is the single most common finding at the

nascent level, and it is frequently underestimated. Formally assigning a named operational owner costs nothing in budget and immediately unblocks progress across several cells at once.

- **When Systems is binding:** The policy rules exist, but the institution cannot extract or verify the necessary records from its software systems. In this situation, the institution is also unlikely to be able to satisfy an external auditor or regulator. This is usually the most resource-intensive axis to upgrade, so planning for systems integration should begin as early as possible.
- **When Process is binding:** Policies, owners, and systems are in place, but they have not been integrated into the everyday workflows of supervision, advising, student milestones, or thesis defense. In other words, good institutional work has been done, but it is not actually reaching graduate students or faculty. The remedy is to insert explicit checkpoints into existing administrative workflows rather than attempting to build brand-new processes from scratch.

Analyze the overall developmental pattern across dimensions rather than relying on an average. Four characteristic institutional patterns tend to recur:

- *Flat at absent or nascent across all five dimensions.* This is a common and completely normal starting baseline. In this situation, you should sequence your improvement efforts in order from Dimension 1 through Dimension 5, focusing initial efforts during the first year on the policy and people axes.
- *Policy running far ahead of everything else.* The institution has published an AI policy, but there are no named operational owners, no retrievable system records, and no integration into academic workflows. The report identifies publishing a single AI policy and declaring the issue resolved as the most common institutional misstep of 2026. The framework addresses this directly: a written policy is only one row in a comprehensive competency grid, not the entire grid.
- *Strong on Dimensions 1 and 2, but weak on Dimension 3.* The institution has clearly articulated the human-judgment demarcation and responsible-use rules, but the software tools researchers actually use every day do not support those standards. This corresponds to the University College London case study discussed in the report's §4.3. In this scenario, the human-in-the-loop discipline cannot be realized at scale until an effective IT procurement gate is put in place.
- *Substantial school and college activity alongside a weak central administration.* This decentralized pattern corresponds to the third case study in the report's §4.3, which the report associates with the US Ivy-Plus and Association of American Universities cluster. In these environments, the central administration may receive a low score, but the institution as a whole is not as weak as that central score suggests. The diagnostic finding is that excellent local practices have not yet been adopted across the wider institution. You should document these school-level instruments in the notes column and focus on scaling them up as a central policy-axis initiative.

Do not compare your profile to another university without confirming the unit of assessment. An assessment evaluating a central university administration and one evaluating a single graduate school are measuring two fundamentally different things.

---

6. Class A–D placement

6.1 What this section is, and what it is not

The Class A–D taxonomy classifies the *substantive scope* of an institution's AI policy documents. This is an entirely separate measurement from the maturity grid, and one should never be derived from the other. For instance, an institution classified as Class C on policy scope can easily sit at the *nascent* level on Dimension 3 if its published research-integrity policy lacks an active IT procurement gate. Conversely, an institution with a Class B policy scope might achieve an *established* rating on Dimension 3 because its central IT and procurement controls are highly mature, even though its written policy covers nothing beyond basic plagiarism rules. Both

situations happen frequently in practice. You should record both measurements independently, and avoid inferring one from the other.

Answer the questions in this section strictly from the written text of the policy documents, not from informal institutional practice. Lived practice is what the maturity grid in §4 is designed to measure.

Declare the unit of assessment. Record clearly whether you are evaluating the central institution or a specific school, college, or graduate school. In the report, the Class D institutions include documents issued directly by doctoral schools rather than central university administrations, so specifying the exact unit of assessment is essential for an accurate evaluation.

6.2 Gate 1: Does a dedicated AI document exist for research and research training?

P1.1. Does an AI-specific policy or guidance document exist that is distinct from the general academic-integrity or student-conduct rules?

> ☐ Yes ☐ Partial ☐ No

Evidence: The document title, URL, issuing body, and date of issue.

P1.2. Does that document apply specifically to doctoral research and research training, rather than only to undergraduate or coursework assessment?

> ☐ Yes ☐ Partial ☐ No

Evidence: The scope clause, quoted directly.

P1.3. Is the document issued centrally or by the graduate school, rather than by a single academic department acting on its own?

> ☐ Yes ☐ Partial ☐ No

Evidence: The issuing authority and the specific institutional population bound by the document.

> ***If P1.1 or P1.2 is No, the placement is Class A.*** *If P1.3 is No or Partial, either redefine the assessment unit to the issuing unit and restart this section, or exclude that document from the declared unit's classification. A department-only document cannot establish the central institution's Class B, C, or D placement.*

6.3 Gate 2: Does the document's scope extend past misconduct into research integrity?

P2.1. Does the policy explicitly address the validity and reproducibility of AI-augmented research, rather than focusing solely on author attribution and plagiarism?

> ☐ Yes ☐ Partial ☐ No

Evidence: The relevant clause, quoted with its exact document reference.

P2.2. Does it address fabrication risks in AI outputs (such as hallucinated citations, synthetic data, or generated images) and mandate verification against primary sources?

> ☐ Yes ☐ Partial ☐ No

Evidence: The relevant clause, quoted directly.

P2.3. Does it define rules for using AI as an analytical validator or checker of the researcher's own work, as distinct from using AI simply as an editorial writing aid?

> ☐ Yes ☐ Partial ☐ No

Evidence: The relevant clause, quoted directly.

P2.4. Does it require explicit disclosure of AI use on research outputs, rather than restricting disclosure rules to assessed coursework?

> ☐ Yes ☐ Partial ☐ No

Evidence: The disclosure clause and the specific research outputs it covers.

> ***If any of P2.1 through P2.4 is answered No or Partial, the institutional placement is Class B.*** *In this case, a dedicated AI policy exists, but its substantive scope is limited to student misconduct and plagiarism framing, extended at most to the standard rule that an AI tool cannot be credited as an author.*

6.4 Gate 3: Does the document extend into competency and examiner discipline?

P3.1. Does the policy state explicit research competency expectations for doctoral students, detailing what a student is expected to understand and do rather than solely listing prohibited behaviors?

> ☐ Yes ☐ Partial ☐ No

Evidence: The clause, quoted directly, along with the specific competencies it defines.

P3.2. Does it establish clear AI competency expectations for doctoral advisors and supervisors?

> ☐ Yes ☐ Partial ☐ No

Evidence: The relevant clause, quoted directly.

P3.3. Does it link to an approved curriculum or a named training pathway designed to develop those competencies?

> ☐ Yes ☐ Partial ☐ No

Evidence: The clause, along with the specific program or pathway it identifies.

P3.4. Does it define explicit examiner-side rules for the doctoral oral examination and thesis declaration, detailing what examiners must review, what they must confirm, and what they are not permitted to do?

> ☐ Yes ☐ Partial ☐ No

Evidence: The examiner clause, quoted directly, and the declaration instrument to which it refers.

> ***If any of P3.1 through P3.4 is answered No or Partial, the institutional placement is Class C.*** *In this case, the policy addresses research integrity, validity, reproducibility, and disclosure, but stops short of defining formal competency standards and examiner verification rules.*
>
> ***If all four questions (P3.1 through P3.4) are answered Yes, the institutional placement is Class D.***

6.5 The published distribution to compare against

The report classified thirty-eight top-tier doctoral universities across fifteen countries and jurisdictions.

Class	Substantive scope	Institutions	Share
A	Generic student-conduct policy with an AI clause appended	4 of 38	approx. 11 percent
B	AI-specific policy whose scope ends at plagiarism and the no-AI-author rule	11 of 38	approx. 29 percent
C	AI-specific policy extending into research integrity and reproducibility	17 of 38	approx. 45 percent
D	AI-specific policy extending into AI literacy, competency, and examiner discipline	6 of 38	approx. 16 percent

Table 6. Distribution and classification of surveyed doctoral universities by policy substantive scope.

Class C is the modal posture across the sample. Taken together, Class C and Class D account for twenty-three of the thirty-eight institutions evaluated (around three in five), while the remaining two in five remain limited to basic plagiarism and student-conduct framing. Approximately one in six institutions in the sample achieves Class D.

Two practical points matter when discussing your own institution's result. First, the operational difference between Class C and Class D is very narrow and specific: it requires adding codified competency expectations for doctoral candidates, advisors, and supervisors, alongside explicit examiner rules for the oral defense and thesis declaration. Because policy class is evaluated strictly from the written document, bridging the gap from Class C to Class D is essentially a drafting task, which makes it one of the most achievable policy updates available. However, actually delivering the training and competencies promised in that policy is an operational question for Dimension 4, which the report's §5.1 identifies as the largest ongoing institutional investment. Second, policy maturity varies widely even within traditional university groupings. For example, the report found that the UK Russell Group spans from Cambridge at Class B to University College London and King's College London at Class D. Comparing your institution to peers based on general academic reputation rather than documented policy scope will give you a misleading picture.

---

7. Crosswalk

Every indicator in this instrument maps to a named section of *Responsible AI in Academic Research: A Competency Framework for Research Training*.

Instrument item	Indicator IDs	Report section
§1.1 Purpose and the unoccupied rung	—	§2.5, §4.2, §4.1
§1.3 Unit of assessment	—	§4.3, §5.2
§2 Evidence rule and the Partial trap	—	§4.1, Appendix A
§3.2 Cumulative ratchet	—	Appendix A (leading cells: "All of *established*, plus...")
§3.3 Dimension = minimum over axes	—	§4.1, §2.5
§3.4 Refusal of a single score	—	§4.3
D1 · policy	D1-policy-N1, E1, E2, L1	§2.1, §4.2, Appendix A row 1, Appendix G
D1 · people	D1-people-N1, E1, L1	§2.1, §2.5 (people axis), §5.3
D1 · systems	D1-systems-N1, E1, L1	§2.1, §2.5 (systems axis), Appendix A row 1
D1 · process	D1-process-N1, E1, E2, L1	§2.1, §2.5 (process axis), §5.3, Appendix A row 1
D2 · policy	D2-policy-N1, E1, E2, L1	§2.2, §3.3, Appendix A row 2
D2 · people	D2-people-N1, E1, L1	§2.2, §2.5 (people axis), §5.3
D2 · systems	D2-systems-N1, E1, L1	§2.2, §2.5 (systems axis)
D2 · process	D2-process-N1, E1, E2, L1	§2.2, §2.5 (process axis), §5.3, §5.4
D3 · policy	D3-policy-N1, E1, E2, L1	§2.3, §4.2, Appendix A row 3, Appendix C
D3 · people	D3-people-N1, E1, L1	§2.3, §2.5 (people axis), §5.1
D3 · systems	D3-systems-N1, E1, E2, L1	§2.3, §2.5 (systems axis), Appendix C
D3 · process	D3-process-N1, E1, L1	§2.3, §2.5 (process axis), §5.4
D4 · policy	D4-policy-N1, E1, L1	§2.4, Appendix A row 4
D4 · people	D4-people-N1, E1, E2, L1	§2.4, §5.3, Appendix A row 4
D4 · systems	D4-systems-N1, E1, L1	§2.4, §2.3 (tutor-mode tooling), Appendix C
D4 · process	D4-process-N1, E1, L1	§2.4, §5.2, §5.3
D5 · policy	D5-policy-N1, E1, E2, L1	§2.5, §4.2, §5.1, Appendix A row 5
D5 · people	D5-people-N1, E1, L1	§2.5, §5.1
D5 · systems	D5-systems-N1, E1, L1	§2.5, Appendix A row 5
D5 · process	D5-process-N1, E1, L1	§2.5, §5.1, §5.4, §5.5
§5.5 Reading the profile	—	§4.3
§6 Class A–D placement, gates 1–3	P1.1–P1.3, P2.1–P2.4, P3.1–P3.4	§1.3

Instrument item	Indicator IDs	Report section
§6.5 Published distribution	—	§1.3, "The finding"

Table 7. Mapping of assessment instrument items and indicator identifiers to framework report sections.

In total, this instrument includes sixty-eight maturity indicators across twenty cells, and eleven policy placement questions across three gates.

---

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