

A4 — Gap-Analysis Worksheet and Action Ladder

Turning a maturity score into named work with named owners: how to read the three sightlines against each other, and sixty priority work packages toward the target rungs of the twenty cells.

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

About this instrument

Purpose	To convert the results of A1, A2, and A3 into a defensible dimensional profile, identify the binding axis in each dimension and the binding dimension overall, and name priority work packages toward the selected target rungs.
Who completes it	The named owner of the institutional AI-readiness benchmarking cycle, working with the Graduate School, the Research Integrity Office, and IT. Part 1 is completed in a half-day workshop once A1, A2, and A3 are in. Part 2 is a reference catalog consulted whenever a cell is selected for improvement.
Time required	Part 1: three to four hours with the completed instruments in the room. Part 2: fifteen minutes per selected rung.
Related report section	Part 4 (§4.1–§4.3), Appendix A, and Part 5 (§5.1–§5.5) of <i>Responsible AI in Academic Research: A Competency Framework for Research Training</i> .
Cells touched	All twenty. Part 1 scores them, and Part 2 supplies three work packages toward the target rungs of each cell.

Table 1. Purpose and structural overview of the assessment instrument.

How to use this document

Part 1 resolves three independent sightlines on the same institution: what leadership can document through instrument A1, what graduate students report in A2, and what advisors and supervisors report in A3. It resolves these into one recorded level per cell, preserving any disagreements rather than averaging them away. In fact, these disagreements are often the most useful diagnostic output.

Part 2 presents an action ladder. In this framework, the gap analysis lists practical institutional improvements in order. Each of the twenty cells has three work packages organized around the target rungs: *absent* → *nascent*, *nascent* → *established*, and *established* → *leading*. Each work package has a named owning role, a required evidence artifact, and an estimated effort band. It supports a later rescore rather than proving the transition by itself.

Terminology. In this document, I use **graduate students** for the people being trained and **advisors and supervisors** for the people who train them. I use **doctoral students** when formal enrolled status is specifically relevant, such as for a progression milestone, an examination rule, or a formal signature block. This term covers everyone enrolled in a research doctorate at any stage, regardless of the local institutional title (*PhD student*, *research student*, *graduate researcher*, or *Doktorand*). **Thesis** refers to the final written work submitted for the degree, which is called the dissertation in much of North America. **Examiners** are the scholars appointed to evaluate it (often

called the dissertation committee in United States practice). Much of the action ladder governs research practice across the entire institution rather than research training alone. Where that is the case, I simply refer to *researchers*.

PART 1 — GAP ANALYSIS

1.1 The three sightlines

A1: Documentary assessment. A1 examines what the institution has officially written down, approved, published, and scheduled for regular review. Its evidence base relies on concrete artifacts: policies, committee minutes, form versions, training registers, procurement records, and formal impact assessments.

A2: Graduate-student evidence. A2 asks graduate students what reached them in practice, what they did, what training they received, and what they know or believe. Those item types have different analytic uses and are kept separate.

A3: Advisor and supervisor evidence. A3 asks advisors and supervisors parallel and supervisor-specific questions from the instructional side: what training they completed, what they require of students, what they verify, and what records they maintain.

None of the three instruments substitutes for the others. A1 establishes what can be supported from institutional records. A2 and A3 provide respondent evidence about selected operational claims and describe the experiences and practices of the people who respond. Part 4 §4.1 of the report defines *established* maturity as having policy, people, systems, and process "all in place and operating at a reviewable cadence". A1 alone cannot demonstrate the operational half of that definition, and a survey response cannot establish that a policy or system exists.

Tagging A2 and A3 items to cells

Every A2 and A3 item carries a cell identifier using the spine's convention: D1-policy, D3-systems, D5-process. The tag records conceptual relevance. It is not an instruction to count that item in a cell score.

Before opening either survey, classify each item as one of four types: **direct operational evidence**, **training exposure**, **knowledge probe**, or **context and diagnostic evidence**. For every direct item that will test an A1 criterion, record the criterion, eligible population, time period, response direction, treatment of Partial, and handling of blank and Not applicable responses. Items that measure prevalence, adverse events, beliefs, confidence, demographics, or open comments remain diagnostic. Do not pool them into a maturity score.

1.2 Computing the level in each cell

Step 1: Prepare the evidence without creating a common survey score

Keep A1 documentary scoring separate from survey analysis. For each predeclared direct operational item, report the full response distribution, eligible number, missing number, and Not applicable number. You may code **Yes = 1.0**, **Partial = 0.5**, and **No = 0.0** to summarize that one directionally aligned criterion. Blank and Not applicable responses are excluded from the item denominator and reported separately. Do not carry that coding across exposure, knowledge, prevalence, adverse-event, confidence, demographic, or open-text items as if they formed one scale.

Dimension 4 requires particular care. Training completion, access, and self-reported confidence are evidence about institutional conditions and exposure. They are not proof of individual

competence unless the institution also uses a direct performance measure aligned with the competency being claimed.

Step 2: Apply the cumulative progression rule (the ratchet) to A1

In this framework, maturity levels build strictly on top of one another. This progression rule is a "ratchet": once a standard is reached, higher levels require maintaining all lower-level requirements. An institution cannot skip intermediate steps or claim a higher maturity level without first meeting every requirement below it.

Each cell's A1 indicators are grouped by maturity level: nascent (N), established (E), and leading (L). The *absent* level has no 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.

Evaluate each level in turn, working upward from nascent:

- **Met:** Every indicator at that rung scores Yes (1.0), and every lower rung is met.
- **In progress:** At least one indicator at that rung scores Yes or Partial, none scores No or is unanswered, and the rung is not met.
- **Not met:** Documentary evidence supports one or more No (0.0) answers.
- **Unresolved:** An unanswered indicator prevents a determination and no documented No already establishes that the rung is not met.

The cell's documented level is the highest rung that is met. For an in-progress or unresolved higher rung, retain the demonstrated level and record the next rung separately. If an unanswered item prevents the nascent rung from being demonstrated, mark the cell **Unresolved/not demonstrated** with no maturity assignment rather than treating the blank as a documented No. This notation does not add a fifth maturity level.

If any axis cell is unresolved 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 another documented cell already fixes the dimension at a lower level independently. Keep this separate from documentary coverage. A coverage verdict of Final does not make an unresolved maturity result final.

This cumulative rule is not optional. Every *leading* cell in Appendix A begins with the phrase "All of *established*, plus...". Consequently, a cell cannot be rated as leading if it fails any established requirement, no matter how impressive its leading evidence might appear.

Step 3: Test each operational criterion against respondent evidence

Test only A1 criteria that make a directly observable claim about operation or reach at the *established* or *leading* level. Select the aligned A2 or A3 item before reviewing results. If several items address different parts of a criterion, report them separately unless the analysis plan defined their combination in advance. Never compute a cell mean from all items carrying the same tag.

The following thresholds are provisional decision rules for a directly aligned criterion. The substantiation rate is the mean of the predeclared Yes = 1.0, Partial = 0.5, and No = 0.0 codes among eligible, nonblank responses. Blank and Not applicable responses are excluded from the denominator and reported separately. These are not empirically calibrated scale cutoffs.

Predeclared substantiation rate	Reading	What it means
70% and above	Survey-consistent	Respondent evidence is broadly consistent with the documented operational claim.

Predeclared substantiation rate	Reading	What it means
40–69%	Mixed	Respondent evidence indicates uneven reach or operation that requires review.
Below 40%	Survey-discrepant	Respondent evidence is inconsistent with a required operational claim.

Table 2. Provisional decision thresholds for criterion-specific respondent evidence.

Always show the observed percentage, full response distribution, eligible number, response rate, and missingness beside the reading. If the predeclared response-rate floor is not met, label the survey evidence **indicative** and do not use it to change a level.

Resolution rules:

- **Survey-consistent:** Retain the documented level for that criterion.
- **Mixed:** Retain the documented level, flag the discrepancy, and record the follow-up needed. Do not lower a level automatically.
- **Survey-discrepant:** Retain the documented level while the discrepancy is adjudicated. Where the item directly measures a required operational condition and coverage is adequate, review the named A1 criterion against all available evidence. Change the maturity result only if that criterion-level adjudication establishes that a specific A1 indicator is or is not satisfied, then run the cumulative A1 scoring rule again. Do not impose an automatic cap from the survey result alone. This review does not apply to a policy-existence or system-existence criterion that respondents cannot directly observe.

Step 4: Resolve A2 against A3

A2 and A3 are anonymous cohort surveys, not matched pairs. Give student evidence priority only for a same-referent, recipient-facing experience, such as whether an AI-use discussion occurred or whether guidance reached the student. A3 commonly reports the proportion of supervisors for whom a practice applies to all or some students, while A2 reports the proportion of individual students describing their own experience. Show those rates side by side. Do not subtract them or call their difference a numerical gap unless a predeclared analysis supplies a common weighted denominator. For policy facts, systems, supervisor behavior not visible to students, or items with different time frames or response scales, report both groups side by side and investigate the difference. Do not declare either group the general ground truth.

Step 5: Handle undocumented practice

When directly aligned A2 or A3 evidence indicates an active practice that the institution cannot substantiate with any written artifact in A1, **retain the A1 documentary level** and flag a practice/codification gap. Respondent evidence does not promote the cell to *nascent*. Achieving the *established* level requires regular operation "at a reviewable cadence" (the report §4.1). An unrecorded practice is not reviewable. It cannot be formally audited, defended to an external regulator, handed over when key personnel change roles, or applied consistently across different schools and colleges.

Step 6: Check coverage

Count the proportion of A1 indicators that received an answer other than blank. **If item coverage falls below 80%, the verdict must be recorded as *Provisional***, and this status must be noted on every page where the assessment appears. A provisional verdict can still be used internally to guide and select improvement work, but it cannot be used to make public claims about the institution's maturity level.

Drafting note on survey response rates. The 80% rule applies only to documentary item coverage in A1. It determines whether that documentary assessment is labeled Final or Provisional. Institutions should additionally set a local floor for A2 and A3 response rates below which respondent evidence is treated as indicative rather than level-changing, for example [25]% of the eligible population with no school or college contributing fewer than [10] responses. Record this response-rate threshold before opening the survey rather than after reviewing the results.

1.3 The binding axis in each dimension

The maturity level for any dimension is determined by the minimum score across its four axes (policy, people, systems, and process). It is not the mean, not the median, nor the mode, and not the highest level achieved on an isolated axis. Part 4 §4.1 of the report requires all four axes to be operating together to achieve the *established* level. A dimension with three axes at *established* and one at *nascent* therefore remains at *nascent*. Advancing along a dimension requires demonstrable achievement across all four axes at that level.

The binding axis is the minimum (the argmin), which is the specific axis holding the overall dimension back. You should report this axis by name whenever you communicate the dimension level. Saying "Dimension 3 is at nascent, with systems as the binding axis" gives leadership actionable information, whereas simply stating "Dimension 3 is at nascent" does not.

When two or more axes tie at the minimum score, **all tied axes are binding** and should be reported. You should sequence subsequent improvement work by practical dependencies rather than by financial cost. The policy rung usually needs to come first, because the corresponding people, systems, and process rungs will inevitably refer to a formal policy artifact that does not yet exist.

1.4 The binding dimension overall

The five dimensions are sequenced in a logical order, where each dimension provides the necessary foundation for the next. If an institution has not made a clear decision on human-in-the-loop oversight in Dimension 1, it has no basis for defining responsible use in practice under Dimension 2. Without an operational definition of responsible use, testing whether research tools promote it under Dimension 3 is essentially meaningless. Similarly, without clear operational rules and approved tooling, the AI literacy curriculum in Dimension 4 has no concrete subject matter to teach. Finally, without these operational foundations across Dimensions 1 through 4, the institutional benchmarking grid in Dimension 5 has nothing reliable to measure.

The binding dimension overall is the lowest-numbered dimension sitting at the lowest maturity level. For example, if Dimension 2 and Dimension 4 are both at the nascent level while all other dimensions are higher, the binding dimension is Dimension 2. This is because any lasting progress on Dimension 4 depends on definitions and standards that must first be established in Dimension 2.

The Dimension 5 interlock

Section 4.1 of the report describes this requirement directly: an institution cannot defensibly claim to be *established* on Dimension 1 while remaining at *nascent* on Dimension 5. This is because having documented evidence of your institutional practices is an essential part of proving that those practices are actively operating rather than merely asserted on paper.

Apply this as a strict consistency check:

Is D5 established? ☐ Yes ☐ No **Resulting assurance status:** ☐ Assured ☐ Unassured

If the answer is **No**, then every claim of *established* or *leading* on Dimensions 1 through 4 must be recorded as **unassured**. Unassured claims may still be discussed in internal planning meetings, but they cannot be published externally, cited in funding applications, or submitted for academic accreditation. To resolve this, the institution must either gather the necessary documentary evidence for Dimension 5 or step its maturity claims down to nascent. This interlock catches a very specific and common institutional problem: having good, working practices in Dimensions 1 and 2 that the university cannot actually prove because no one maintains the formal records.

1.5 What the instrument does not produce

There is no single institution-wide score, and this instrument will not compute one. As the report emphasizes, "the diagnostic value of the grid is the dimensional pattern, not a single score" (§4.3). Reducing an entire university's readiness to a single uninformative number erases the dimensional profile and obscures the exact operational bottlenecks the assessment was designed to uncover.

The valid outputs of this assessment are:

1. The twenty recorded cell maturity levels.
2. The five overall dimension levels, with the binding axis clearly identified for each.
3. The single binding dimension for the institution overall.
4. The assurance status determined by the Dimension 5 interlock.
5. The A1 documentary coverage percentage.
6. The documentary coverage status (*Final* or *Provisional*).

Validation status. This is a criterion-referenced, source-anchored operational instrument. It has not been field-calibrated as a psychometric scale. *Final* means that A1 documentary coverage reached the stated threshold. It does not mean that every survey inference has been independently validated. The maturity categories are ordinal and the distance between adjacent levels should not be treated as equal. Longitudinal comparisons require the same assessment unit, instrument version, survey populations, fielding window, and coding rules. The pack should not be used to rank institutions or interpret score differences as precise quantities.

Determining an institution's policy class (Classes A through D) is an entirely separate exercise. The four policy classes described in Part 1 §1.3 of the report define the *scope* of an institution's written AI policy: assessment-only (Class A), misconduct-only (Class B), research-integrity (Class C), or comprehensive AI-literacy and examiner-discipline (Class D). Policy class is not calculated from maturity levels, and the relationship does not run in reverse either. A Class C institution, for example, can sit at any maturity level across the twenty grid cells.

1.6 Worked interpretation: reading a mismatched pattern

Pattern A: Policy established, people nascent

What you are looking at. The policy document exists, it was formally approved, and it is published on the university website. However, the people responsible for carrying it out have not been named, trained, or allocated dedicated workload time. When asked, graduate students and supervisors cannot describe the policy's requirements in their own words.

What it predicts. Compliance will tend to concentrate only in those schools and colleges where individual leaders happen to take a personal interest, and central leadership will have no clear

visibility into where those pockets are. When the first serious incident occurs (such as a fabricated citation reaching an external examiner or a researcher uploading unpublished data into a commercial consumer tool), the case will be judged against a policy that staff and students were never taught. In post-incident reviews, institutions often describe this as "clear guidance that was not followed", which inappropriately shifts institutional responsibility onto an individual.

What to do. Do not write more policy. Policy drafting is something universities already do well, and producing another document creates an illusion of progress without changing daily practice. Instead, select the people rung in that cell, alongside the process rung that embeds the requirement into an obligatory workflow. Policy only influences real practice when it is tied to mandatory training and required forms, not merely by posting it on a website.

Pattern B: People established, policy nascent

What you are looking at. Genuine, good-faith practice that runs purely on the personal initiative of dedicated individuals. Graduate students report having substantive discussions about responsible AI use, and supervisors describe thorough verification routines, but no approved university policy formally requires or supports any of it.

What it predicts. This creates three predictable vulnerabilities. First, the practice disappears as soon as key staff members change roles or leave the university. Second, it remains invisible across the rest of the institution, meaning external examiners apply inconsistent standards and graduate students in neighboring departments face completely different requirements. Third, because the practice is uncoded, the university cannot demonstrate compliance to external research funders or regulatory bodies.

What to do. Codify what is already working rather than trying to design a new system from scratch. Document the existing grassroots practice, formally designate the staff who run it as its official owners, and submit the policy to [COMMITTEE] for approval. This is typically a six-to-ten-week task rather than a six-month project, and it almost always results in a more practical document than a committee starting from a blank page.

Pattern C: Systems established, process nascent

What you are looking at. The institution has purchased and deployed the necessary infrastructure (such as enterprise software licenses, structured declaration fields, and records systems), but actual workflows have not changed. As a result, the new fields remain completely empty or are filled with "N/A".

What to do. Select the process rung. If completing a data field is not required as part of an obligatory operational step, people will not fill it out. You must attach the requirement to a mandatory administrative gate, such as a progression milestone that cannot be signed off, a thesis submission that cannot be accepted, or an examiner appointment that cannot be finalized without it.

Pattern D: Process established, systems nascent

What you are looking at. Administrative workflows require all the right disclosures and checks, but the institution's information systems do not support them. Declarations arrive as scattered email attachments, free-text paragraphs, or separate PDF files, and an administrator is forced to track everything manually in a spreadsheet.

What to do. Select the systems rung. This manual pattern can feel deceptively stable, but it is quietly expensive. While it gets the job done in the short term, it produces no standardized, reportable data and consumes professional staff time indefinitely. It is also the workflow pattern that fails most rapidly when student or submission volumes increase.

1.7 The triangulation table template

First, assign a reconciliation record ID and complete one row for every A1 criterion tested with respondent evidence. Duplicate the blank row as needed. Do not combine criteria that have different populations, referents, time periods, directions, eligibility rules, or thresholds.

Reconciliation record ID	A1 criterion ID	Cell and rung	Survey item / key	Population, referent, and time period	Predeclared threshold
--------------------------	-----------------	---------------	-------------------	---------------------------------------	-----------------------

Table 3A. Criterion-level test specification for respondent evidence used to examine an A1 operational claim.

Use the same reconciliation record ID to complete its evidence and adjudication record.

Reconciliation record ID	Full distribution and eligible N	Reading	Criterion adjudication	Follow-up
		consistent / mixed / discrepant / indicative	retained / A1 indicator changed after review	

Table 3B. Criterion-level evidence and adjudication record linked to Table 3A by reconciliation record ID.

After every tested criterion has been adjudicated, complete one roll-up row per cell. Twenty rows. Where a cell contains several tested criteria, summarize their readings without averaging them and attach the criterion-level records above.

Cell	A1 documented level	A1 coverage	A2 criterion evidence	A3 criterion evidence	Survey reading
D1-policy		%	see criterion records	see criterion records	consistent / mixed / discrepant / indicative / not tested
D1-people					
D1-systems					
D1-process					
D2-policy					
D2-people					
D2-systems					
D2-process					
D3-policy					
D3-people					

Cell	A1 documented level	A1 coverage	A2 criterion evidence	A3 criterion evidence	Survey reading
D3- systems					
D3- process					
D4- policy					
D4- people					
D4- systems					
D4- process					
D5- policy					
D5- people					
D5- systems					
D5- process					

Table 4A. Cell-level evidence roll-up recording documentary levels and criterion-specific survey readings.

Use the same Cell code to record the decision and action fields for each roll-up row.

Cell	Recorded level	Gap type	Binding axis?	Cell owner (role)	Rung selected
D1- policy		paper / practice / transmission / aligned / documentary-only			
D1- people					
D1- systems					
D1- process					
D2- policy					
D2- people					
D2- systems					
D2- process					

Cell	Recorded level	Gap type	Binding axis?	Cell owner (role)	Rung selected
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 4B. Cell-level decisions and work-package selections linked to Table 4A by Cell code.

Gap types defined:

- *Paper gap*: The documented level in A1 is higher than directly aligned survey evidence supports. The artifact exists, but a required operational condition appears not to reach the relevant population consistently.
- *Practice gap*: Directly aligned A2 or A3 evidence indicates an active practice that the institution cannot substantiate with written documentation. The practice may be operating, but formal institutional codification is missing.
- *Transmission gap*: Advisors and supervisors in A3 report that a recipient-facing practice occurs, but graduate students in A2 report that it does not. This indicates that implementation may be breaking down between supervisors and students and should be investigated.
- *Aligned*: The three sightlines agree on the level of practice, regardless of whether that level is absent, nascent, established, or leading.
- *Documentary-only / not tested*: No directly aligned respondent item tests the cell or criterion. Retain the A1 documentary result and do not describe the untested survey sightlines as aligned.

1.8 The dimensional profile (the reportable output)

Dimension	Level	Binding axis	Gap type on the binding axis	Rung selected this cycle	Owner
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					

Table 5. Dimensional profile reporting maturity levels, binding axes, gap types, selected rungs, and owners.

Binding dimension overall: ____ Is D5 established? ☐ Yes ☐ No **Assurance status:** ☐ Assured ☐ Unassured **Verdict status:** ☐ Final ☐ Provisional (A1 coverage below 80%)

1.9 Readiness-to-report checklist

☐ Yes ☐ Partial ☐ No: Every A1 indicator has an answer or a recorded written explanation for why it was left blank.

☐ Yes ☐ Partial ☐ No: Every A2 and A3 item has a recorded analytic use, response direction where applicable, missing-data rule, and conceptual cell mapping.

☐ Yes ☐ Partial ☐ No: The cumulative progression rule (the ratchet) was strictly applied, so that no cell is recorded at the leading level while any established requirement remains unmet.

☐ Yes ☐ Partial ☐ No: Every dimension maturity level equals the minimum score across its four constituent axes.

☐ Yes ☐ Partial ☐ No: Every mixed or survey-discrepant criterion includes a formal written resolution note in the triangulation table.

☐ Yes ☐ No: The Dimension 5 interlock has been checked for every claim at the established level or above.

☐ Yes ☐ Partial ☐ No: No single, aggregated institution-wide score appears anywhere in the reported output.

☐ Yes ☐ Partial ☐ No: Each of the twenty cells has an accountable institutional role assigned as its owner.

Completed by: [NAME], [ROLE] · **Date:** [DATE] · **Received by:** [COMMITTEE] on [DATE]

PART 2 — THE SIXTY-RUNG ACTION LADDER

How to read a rung

Each of the twenty cells on the grid has three action entries organized around the three target transitions (*absent* → *nascent*, *nascent* → *established*, and *established* → *leading*). Each listed action is a priority work package toward its target rung. It is not, by itself, proof that the cell or dimension has advanced. Each entry specifies four concrete elements:

- **The action:** A direct, practical instruction that a designated person can begin this quarter. Examples include "Draft X and submit it to [COMMITTEE] for formal approval", "Designate an accountable institutional role for Z", or "Establish the W register and define its annual review schedule". The ladder avoids vague aspirations like "improve awareness" or "embed a culture of compliance".
- **The owner:** A specific institutional role rather than an individual person. This role is formally accountable for ensuring the required artifact exists and is maintained, though they may delegate drafting tasks.
- **The artifact:** The verifiable document or system record produced by the work package. It must satisfy the same standard that instrument A1 requires: an external, skeptical auditor should be able to inspect it directly.
- **The effort band:** The estimated calendar time (in weeks or months) required to produce and approve the artifact under standard institutional governance cycles. In practice, committee meeting schedules, rather than drafting time, dictate the duration. If your relevant approval committee meets only twice a year, you should adjust these timelines accordingly.

Rung identifiers combine the cell ID with the target maturity level: D1-policy → nascent, D3-systems → established, or D5-process → leading.

Work through one rung at a time within each cell. Under the cumulative progression rule, an *established* rung assumes that the corresponding *nascent* foundation is already in place, and a *leading* rung assumes that the *established* foundation is operating. After completing a work package, rescore every applicable A1 indicator at the target rung and every rung below it. Advance the cell only when all of those indicators are Yes. Record a partially completed rung as supporting work, not as level-completing evidence. A cell currently at the absent level cannot jump directly to established in a single leap, because the established action has no underlying policy or system foundation to build upon.

These work packages are a planning catalogue, not a one-to-one crosswalk to the A1 criteria. A heading names the cell and planning direction that the package supports. Before commissioning it, list the exact A1 indicator IDs that the institution intends to rescore. If the completed package does not satisfy those criteria, record it as preliminary supporting work and do not change the maturity level.

A work package supplies evidence only toward its named cell. A D5-people work package, for example, can support movement in D5-people only. Dimension 5 moves to a target level only when its policy, people, systems, and process cells all qualify at that level.

Each cell description begins with an italicized summary mapping Appendix A's requirements for that dimension to the specific axis. In the companion report, Appendix A is structured by dimension and level. The four-axis breakdown (policy, people, systems, and process) follows the analytical framework set out in the report §2.5.

Role glossary

The ladder uses a consistent set of institutional role names. Map them to your university's specific titles before circulating this document.

Role used in this ladder	Local equivalents
Deputy Vice-Chancellor (Research)	Pro-Vice-Chancellor (Research), Vice-President for Research, Vice-Rector for Research, Deputy Provost for Research
Dean of the Graduate School	Dean of Graduate Studies, Director of the Doctoral School, Head of the Graduate Academy, Associate Dean (Research Training)
Director of Research Integrity	Research Integrity Adviser or Officer, Head of Research Ethics and Integrity
Chair, Graduate Studies Committee	Chair of the Graduate Research Committee, Doctoral Board, Board of Graduate Studies, Graduate Council
[COMMITTEE]	Academic Board, Senate, Research Committee, University Council (whichever body approves policy of this class)
Chief Information Officer	Director of Information Technology, Chief Technology Officer
Chief Information Security Officer	Head of Information Security
Data Protection Officer	Privacy Officer, <i>Datenschutzbeauftragte(r)</i> , Head of Information Governance
University Librarian	Director of Library and Information Services
Associate Dean (Research)	Vice-Dean (Research), Associate Dean for Research, Director of Research for a school or college
Director of Research Information Systems	Head of Research Systems, Research Office systems lead

Table 6. Standard ladder roles and corresponding local institutional title equivalents.

Dimension 1 — Human-in-the-loop discipline

D1-policy

Appendix A, read on this axis: at nascent, a stated institutional position. At established, a task-level demarcation published as policy. At leading, published scoring reviewed at a stated cadence.

D1-policy → nascent — Publish the position. Draft a one-page institutional position statement establishing that AI-generated content must be verified before use and cannot replace human expert judgment. Clarify that this requirement applies across all research and research training, not just student assessment. Submit the statement to [COMMITTEE] for formal approval. - *Owner:* Deputy Vice-Chancellor (Research), drafted by the Director of Research Integrity. - *Artifact:* Approved position statement carrying a version number, approval date, next review date, and a public URL. - *Effort:* 4–8 weeks.

D1-policy → established — Publish the labor-versus-judgment schedule. Draft a research-task demarcation schedule as a numbered appendix to the AI-in-research policy. This schedule should clearly distinguish which tasks AI may augment under human verification, which tasks AI may not substitute for, and which tasks remain strictly human-only. Use Appendix G of the report as a starting template, and require each school or college to develop discipline-specific extensions within one semester. - *Owner:* Chair, Research Integrity Committee, with the Dean of the Graduate

School. - *Artifact*: Policy clause containing the task schedule, the committee minute approving it, the school and college extensions, and the defined review cadence. - *Effort*: 3–5 months.

D1-policy → leading — Publish the score and the method. Publicly report the university's own Dimension 1 maturity level alongside its supporting evidence and scheduled re-assessment date, in an open location that external readers can easily find. Embed the regular review cadence directly into the policy text so that it survives changes in leadership personnel. - *Owner*: Deputy Vice-Chancellor (Research). - *Artifact*: Public scoring page showing maturity level, evaluation method, supporting evidence, assessment date, and next review date, alongside a policy clause establishing the review cadence. - *Effort*: 2–3 months after reaching established.

D1-people

Appendix A, read on this axis: at nascent, a named position-holder. At established, supervision that enforces the demarcation and examiner training that references it. At leading, published completion evidence.

D1-people → nascent — Name the owner in the policy. Add a formal accountability table to the position statement specifying the institutional role accountable for the labor-versus-judgment demarcation, the role responsible for maintaining it, and an official delegate. List institutional roles rather than individual names. - *Owner*: Deputy Vice-Chancellor (Research). - *Artifact*: Accountability table published within the policy, specifying role, delegate, and review date. - *Effort*: 1–2 weeks.

D1-people → established — Make the demarcation part of supervisor registration. Incorporate the labor-versus-judgment demarcation into supervisor induction programs and mandatory refresher courses. Require all principal supervisors to complete this training once per [three]-year registration cycle, and make completion a mandatory condition for remaining on the register of approved supervisors. - *Owner*: Dean of the Graduate School, working with Associate Deans (Research). - *Artifact*: Training register documenting completion rates by school and college against the active supervisor list, with the registration requirement incorporated into the supervision policy. - *Effort*: 6–9 months to complete the first cycle.

D1-people → leading — Publish completion by school and college and act on the outliers. Submit annual supervisor and examiner completion rates to [COMMITTEE], publish these figures across the university, and require any school or college falling below the established minimum threshold to submit a formal remediation plan with concrete deadlines. - *Owner*: Dean of the Graduate School. - *Artifact*: Annual completion report detailing figures for each school and college, the committee minute receiving it, and any submitted remediation plans. - *Effort*: 3 months to publish the first report, then annually.

D1-systems

Appendix A, read on this axis: at nascent, the thesis or enrollment system records AI use. At established, the field uses the published task categories. At leading, outcome data is extracted and reported.

D1-systems → nascent — Add the AI-use field to the student system. Add a dedicated field to the thesis-submission or doctoral-enrollment system in which each student records whether generative AI was used. Use a controlled field rather than a note stored in an unrelated document. - *Owner*: Director of Student Systems, with the Graduate School as business owner. - *Artifact*: Released form or system screen, the field definition, and the date it entered production. - *Effort*: 6–10 weeks.

D1-systems → established — Put the demarcation into the submission and examination forms. Embed a structured AI-use declaration into the thesis submission workflow and the examiner report form within the student records system. Use the policy's demarcation categories as standardized selection fields rather than relying on open text boxes, since free-text responses cannot be tracked reliably and are rarely audited. - *Owner*: Director of Student Systems, with the

Graduate School as business owner. - *Artifact*: Released system form version, the data dictionary defining each field, and the initial data extract of completed declarations. - *Effort*: 4–6 months.

D1-systems → leading — Stand up the outcome extract. Establish a standard quarterly reporting extract that tracks thesis declaration completion rates, examiner flag frequencies, and supervision audit findings. Deliver these reports to [COMMITTEE] each quarter and publish annual aggregate summaries. - *Owner*: Director of Research Information Systems, working with the Graduate School. - *Artifact*: Written data report specification, four completed quarterly extracts, and the published annual summary figures. - *Effort*: 2–4 months to define the specification, then ongoing.

D1-process

Appendix A, read on this axis: at nascent, the position touches at least one milestone. At established, enforcement through supervision and checking at examination. At leading, an audit of the practice.

D1-process → nascent — Add the question to the first formal progress review. Add a mandatory question to the template for the student's first formal progress review (such as the first-year probationary review, or advancement to candidacy in United States practice). Require students to describe which research tasks in their project will involve AI assistance and which analytical judgments remain strictly their own. This requires a substantive written response rather than a simple checkbox. - *Owner*: Dean of the Graduate School. - *Artifact*: Revised milestone review template effective from [DATE], featuring the mandatory question alongside a file of completed forms from the initial cohort. - *Effort*: 4–8 weeks.

D1-process → established — Write the defend-your-use rule into the examination procedure. Update the formal examination regulations to require doctoral candidates to be capable of explaining and defending their AI use, as well as the complete contents of their thesis, during the oral examination. Embed this requirement directly in the formal procedure rather than in informal advisory guidance, and attach an examiner briefing note to every formal appointment letter. - *Owner*: Chair, Graduate Studies Committee. - *Artifact*: Amended examination procedure clause with its formal approval minute, and the examiner briefing note attached to standard appointment letters. - *Effort*: 3–6 months.

D1-process → leading — Audit completed examinations annually. Conduct an annual audit on a randomized sample of completed examinations to check whether AI declarations were completed, whether examiners reviewed them, and what issues emerged. Report these findings to [COMMITTEE] and publish the results. - *Owner*: Chair, Graduate Studies Committee, working with the Director of Research Integrity. - *Artifact*: Annual audit report detailing sample size, methodology, findings, and any remediation actions taken in response. - *Effort*: 3 months to complete the first audit, then annually.

Dimension 2 — Responsible use in practice

D2-policy

Appendix A, read on this axis: at nascent, a single published disclosure template. At established, task-level rules for all four AI-use modes. At leading, annual compliance evidence and a scheduled review against publisher and funder movement.

D2-policy → nascent — Publish one disclosure template. Publish a single institutional AI-use disclosure template that specifies exactly what researchers must document (tool name, version, date of use, specific research step, and purpose) and where this statement must appear. This should be a single, uniform template for the entire institution, because departmental or school-level variations create the very inconsistencies the template is meant to eliminate. - *Owner*: Chair, Research Integrity Committee. - *Artifact*: Published institutional template carrying a version

number, alongside the policy clause requiring its use. - *Effort*: 4–8 weeks. (*Companion instrument P2 provides sample standards and template statements.*)

D2-policy → established — Write the four mode rules. Develop task-level operational rules for each of the four primary AI-use modes as separate, numbered policy sections: search (every cited source must be verified against the original primary publication), co-author (AI assistance is permitted only on text originally drafted by the researcher, with no substitution for text they cannot independently explain), validator (AI critiques serve as initial prompts rather than completed peer reviews, and must be tested across different model families), and tutor (prioritize corpus-grounded tutoring tools, and require demonstrable independent recall before treating AI assistance as completed learning). Submit all four sections to [COMMITTEE] together. - *Owner*: Chair, Research Integrity Committee, working with the Dean of the Graduate School. - *Artifact*: Approved policy containing the four distinct mode sections, the formal approval minute, and the scheduled review date. - *Effort*: 4–6 months.

D2-policy → leading — Publish annual compliance evidence and schedule the review. Publish an annual report on policy compliance (covering disclosure completion rates, results from citation-verification audits, and examiner-flag trends), and establish a standing annual review to update the four mode rules in response to changes in publisher guidelines and funding agency policies. - *Owner*: Director of Research Integrity. - *Artifact*: Published annual compliance report and a dated review record detailing updates made to the rules and their rationale. - *Effort*: 3 months to publish the first report, then annually.

D2-people

Appendix A, read on this axis: at nascent, someone owns the template. At established, the rules are taught and the supervision conversation happens. At leading, participation is measured and published.

D2-people → nascent — Name the owner and publish the question route. Designate a specific institutional role as the owner of the disclosure template, and publish a clear channel where researchers can submit procedural questions, backed by a committed turnaround time. When researchers have no clear place to ask questions, departments inevitably resort to inconsistent local improvisations. - *Owner*: Deputy Vice-Chancellor (Research), delegating operational management to the Director of Research Integrity. - *Artifact*: Designated role recorded in the policy accountability table, published contact point, and formal service-level response commitment. - *Effort*: 1–2 weeks.

D2-people → established — Require the supervisor-student AI-use agreement. Deliver training on the four AI-use mode rules to advisors, supervisors, and examiners. Require every incoming doctoral student and their primary supervisor to complete and submit a written AI-use agreement within the first [three] months of enrollment. - *Owner*: Dean of the Graduate School. - *Artifact*: Supervisor and examiner training register, alongside on-file submitted agreements for the incoming cohort, tracked by school and college. (*Companion instrument G1 provides the agreement template.*) - *Effort*: 6–9 months.

D2-people → leading — Publish agreement and training rates by school and college. Publish annual completion rates for AI-use agreements and training programs across every school and college. Require any department falling below the university's established benchmark to submit a remediation plan with concrete actions and target dates. - *Owner*: Dean of the Graduate School. - *Artifact*: Published departmental completion figures with year-on-year comparisons, alongside all submitted remediation plans. - *Effort*: 3 months to publish the first report, then annually.

D2-systems

Appendix A, read on this axis: at nascent, the template is where researchers work. At established, disclosure is structured data aligned to publisher convention. At leading, the aggregate is published and contributed.

D2-systems → nascent — Put the template into the systems people already use. Embed the official disclosure template directly into institutional thesis templates, manuscript submission checklists, and human research ethics applications, ensuring each includes an active version stamp. Templates that reside only on an obscure policy webpage will rarely reach the researchers who need them most. - *Owner:* University Librarian, working with the Research Office. - *Artifact:* Version-stamped disclosure templates embedded across all named systems, with designated operational owners for each integration. - *Effort:* 4–8 weeks.

D2-systems → established — Make disclosure a structured field. Convert disclosure records from unstructured free text into standardized data fields (capturing tool name, version, date, use mode such as search, co-author, validator, or tutor, and specific research task) within the thesis repository and institutional research publications database, aligned with major academic publisher standards. - *Owner:* Director of Research Information Systems. - *Artifact:* Field schema, data dictionary, and an initial data extract demonstrating completed structured disclosures. - *Effort:* 4–6 months.

D2-systems → leading — Publish the aggregate and supply it to the sector. Publish annual aggregated disclosure metrics (including total volume, distribution across use modes, and completion rates), and submit this anonymized dataset to a recognized higher-education benchmarking organization. - *Owner:* Director of Research Information Systems, working with the Deputy Vice-Chancellor (Research). - *Artifact:* Published summary dataset with a defined schema, alongside formal records of data submissions to external benchmarking bodies. - *Effort:* 2–3 months.

D2-process

Appendix A, read on this axis: at nascent, disclosure is acknowledged. At established, the verification expectation is operationalised in supervision and at examination. At leading, verification is audited.

D2-process → nascent — Put disclosure on the submission checklist. Add a mandatory AI disclosure check to the thesis submission process so that no thesis can be accepted for examination without a completed response, including an explicit declaration if no AI tools were used. - *Owner:* Manager, Graduate Research Examinations. - *Artifact:* Updated submission checklist featuring the mandatory prompt, implementation date, and completed checklists from the first month of operation. - *Effort:* 3–6 weeks.

D2-process → established — Require verification evidence at the mid-program review. Update the mid-program progress review requirements so that doctoral candidates formally attest, and their supervisors countersign, that every citation in the draft thesis derived via AI search has been verified against the original primary publication. Make this signed attestation a mandatory prerequisite for passing the milestone. - *Owner:* Chair, Graduate Studies Committee. - *Artifact:* Revised progress review form containing the attestation and countersignature blocks, accompanied by review panel guidance on handling unverified citations. - *Effort:* 3–5 months.

D2-process → leading — Run the citation-verification audit. Conduct an annual audit on a randomized sample of submitted theses, verifying a fixed sample of citations per thesis against original publications, and publish the observed rates of unresolvable or misattributed sources. - *Owner:* Director of Research Integrity, working with the University Librarian. - *Artifact:* Written audit protocol, published error rates, and documentation of corrective actions implemented. - *Effort:* 3–4 months to conduct the first audit, then annually.

Dimension 3 — Tooling that promotes responsible use

D3-policy

Appendix A, read on this axis: at nascent, a gate exists though its criteria are vendor-led. At established, a published procurement standard carrying the six observable properties, with a documented exception process. At leading, a published audit and a documented improvement cycle.

D3-policy → nascent — Issue the interim data rule. Draft and approve a concise interim policy stating that unpublished research data, participant information, and draft manuscripts may only be processed using enterprise-contracted AI services that have training-on-user-input explicitly disabled. Publish a register of currently approved institutional deployments. The March 2023 Samsung data leaks (where three proprietary data exposures occurred within twenty days, prompting an immediate corporate ban on consumer AI tools) provide the report's canonical demonstration that managing data security is an institutional procurement responsibility rather than an individual researcher's task (the report §2.3). - *Owner:* Chief Information Security Officer, working with the Deputy Vice-Chancellor (Research). - *Artifact:* Approved interim data policy with effective date, alongside a published register of approved AI deployments. - *Effort:* 3–6 weeks.

D3-policy → established — Publish the procurement standard. Publish an institutional procurement standard that evaluates all prospective AI software against six observable criteria: (1) verifiable citation handling, (2) robust data residency and training-on-input protections, (3) explicit reporting of model uncertainty, (4) reproducibility tied to specific model versions, (5) full auditability via session logs, and (6) local or open-source hosting options where feasible. Include a standardized evaluation scorecard and a formal exception workflow for specialized research projects that require consumer tools. - *Owner:* Chief Information Officer, working with the Chief Procurement Officer. - *Artifact:* Published standard detailing all six criteria, the scoring rubric, and an active exception register containing worked entries. - *Effort:* 4–7 months.

D3-policy → leading — Publish the audit and adopt a management-system discipline. Publish periodic procurement audit reports on a defined schedule, and manage the university's AI software portfolio under an auditable management framework aligned to recognized international standards, supported by independent external compliance assessments. - *Owner:* Chief Information Officer. - *Artifact:* Published procurement audit report, external assessment certification, and an improvement action log with designated owners and target dates. - *Effort:* 9–18 months.

D3-people

Appendix A, read on this axis: at nascent, someone can say yes or no. At established, a standing panel with the right seats applies the standard. At leading, vendor claims are independently verified.

D3-people → nascent — Name the officer who signs off research AI tools. Formally record in the university delegation schedule the designated officer who must authorize any AI tool that processes institutional research data. Publish a clear submission procedure and commit to a defined response timeframe. When no clear approval pathway exists, individual labs make informal software choices and leadership only discovers them after a security or integrity incident. - *Owner:* Deputy Vice-Chancellor (Research), working with the Chief Information Officer. - *Artifact:* Delegation schedule amendment, published request workflow, and service-level agreement for response times. - *Effort:* 2–4 weeks.

D3-people → established — Stand up the AI tooling review panel. Establish a standing AI tooling review panel with designated representatives from research integrity, information security, data protection, library and information services, and two research-active faculty members from different academic disciplines. Provide the panel with formal terms of reference, a published meeting schedule, and the explicit authority to reject non-compliant tools. - *Owner:* Chief Information Officer as chair, with research representatives appointed by the Deputy Vice-

Chancellor (Research). - *Artifact*: Approved terms of reference, membership roster by role, published meeting calendar, and minutes from the first three meetings. - *Effort*: 3–4 months.

D3-people → leading — Commission independent verification of vendor claims. Engage an independent third party with no commercial ties to software vendors to audit vendor claims on a representative sample of deployed AI tools, and publish the institutional management response. For example, the 2024 Stanford RegLab study found hallucination rates between 17 and 33 percent across three retrieval-augmented legal research tools whose vendors claimed retrieval mechanisms eliminated false citations (the report §2.3). The university's own empirical testing provides the definitive evidence. - *Owner*: Chief Information Officer, working with the Chief Audit Executive. - *Artifact*: Commissioned audit report naming tested tools, evaluation methodology, and observed findings, accompanied by a formal management response. - *Effort*: 4–6 months to complete the initial review, then on a regular schedule.

D3-systems

Appendix A, read on this axis: at nascent, the approved-tool register is maintained. At established, protected enterprise deployments and published session-log retention are in place. At leading, deployed tools are retested for citation and source resolvability on a stated cadence.

D3-systems → nascent — Build the tool register. Create and maintain the official register of AI tools approved for research use. Validate it against procurement records, IT sign-in logs, and departmental returns so unapproved tools are identified for a separate remediation decision rather than silently treated as approved. For every approved entry, record the software category, designated owner, approval decision and date, license tier (consumer versus enterprise contract), data classification levels permitted, and geographic data storage location. - *Owner*: Chief Information Officer, working with Associate Deans (Research). - *Artifact*: Published approved-tool register containing the required attributes and approval record for every entry, with a defined update schedule. - *Effort*: 6–10 weeks.

D3-systems → established — Provide protected deployments and retain their logs. Make enterprise-tenanted deployments with training-on-input disabled available for projects involving unpublished or confidential data. Record data residency and training-on-input status in the tool register, and publish a records-governance schedule that states the interaction-log retention window for every covered deployment. - *Owner*: Chief Information Officer, working with the Data Protection Officer and records manager. - *Artifact*: Tenancy configuration records, complete register entries for residency and training-on-input status, and the published retention schedule. - *Effort*: 4–6 months.

D3-systems → leading — Retest deployed tools on a stated cadence. Run standard citation and source-resolvability tests against each deployed tool on a published schedule, recording the model version, test set, empirical result, and next test date. Include retrieval-augmented tools because they may replace invented citations with subtler misattributions of real sources (the report §2.3). - *Owner*: University Librarian, working with the Chief Information Officer. - *Artifact*: Dated test records for every deployed tool and the approved recurring test schedule. - *Effort*: 3 months for the first cycle, then recurring on cadence.

D3-process

Appendix A, read on this axis: at nascent, the tool question reaches the research workflow. At established, the impact assessments the jurisdictions require are complete and retrievable. At leading, a documented improvement cycle tracks model-version change.

D3-process → nascent — Put the AI question in the data-management plan. Add two mandatory questions to the institutional data management plan template: which AI tools will process this research data, and through which enterprise-approved deployment. Ensure data management plans cannot receive approval without completing these items. - *Owner*: Director of Research

Services. - *Artifact*: Updated data management plan template effective from [DATE], featuring the mandatory questions, alongside completed sample plans from active research projects. - *Effort*: 4–6 weeks.

D3-process → established — Complete and register the impact assessments. Complete the formal impact assessments required across all jurisdictions where the institution operates, and maintain them in a single, retrievable register. This includes: (1) a Fundamental Rights Impact Assessment under Article 27 of the EU AI Act for Annex III §3 high-risk uses (such as admissions, learning-outcome evaluation, education-level assignment within institutions, and proctoring) where the institution is within Article 2's territorial scope and is a body governed by public law or a private entity providing public services, effective August 2, 2026, (2) a Data Protection Impact Assessment wherever personal data is processed, and (3) equivalent regulatory instruments for any other jurisdiction in which the institution operates. - *Owner*: Data Protection Officer, working with the Chief Information Officer. - *Artifact*: Completed and dated impact assessments, a central index register identifying the owner of each, and defined document retention rules. - *Effort*: 6–12 months.

Drafting note on jurisdictions. Institutions operating across multiple countries will need to maintain several regulatory instruments covering overlapping areas. Identify all relevant jurisdictions where the university operates or recruits students before beginning, and clearly identify the specific assessment tool required for each. Do not assume that completing an assessment in one jurisdiction satisfies requirements in another.

D3-process → leading — Operate the improvement cycle. Re-evaluate deployed tools whenever underlying model versions change, decommission tools that fail re-testing, and maintain a documented record of all retention and retirement decisions. This monitoring cycle exists specifically to catch regressions caused by model version updates. - *Owner*: Chief Information Officer. - *Artifact*: System change log connecting each re-test to a model version update, detailing decisions, evidence, and effective dates. - *Effort*: 3 months to establish, then ongoing.

Dimension 4 — AI-literate humans

D4-policy

Appendix A, read on this axis: at nascent, provision exists but is unintegrated. At established, the six competencies are named curriculum outcomes. At leading, evidence on provision, reach, and selected aligned performance outcomes is published and externally reviewed.

D4-policy → nascent — Name the six competencies in the doctoral policy. Incorporate a formal statement into the doctoral policy establishing that AI literacy is a core research-training requirement rather than an optional software skill, and explicitly define the six core competencies: (1) citation verification, (2) model and parameter specification, (3) prompt-as-fork-in-the-garden discipline, (4) model heterogeneity in adversarial review, (5) sycophancy detection and human-as-verifier discipline, and (6) structured failure-mode reporting. - *Owner*: Dean of the Graduate School. - *Artifact*: Approved policy amendment detailing all six competencies, with its formal approval date. - *Effort*: 6–10 weeks.

D4-policy → established — Make them approved curriculum outcomes. Submit the six competencies through standard academic governance channels as formal, approved learning outcomes within doctoral research methods training, with each mapped to specific assessment methods. Formal curriculum approval is the crucial step that turns aspirational policy into required educational delivery. - *Owner*: Chair of the relevant curriculum committee, on a formal proposal from the Dean of the Graduate School. - *Artifact*: Approved course and module outlines

showing the six competencies mapped to specific assessment tasks, accompanied by the formal approval minute. - *Effort*: 6–9 months, aligned with the institutional curriculum approval calendar.

D4-policy → leading — Publish outcome evidence and commission an external review. Publish evidence on provision, reach, and selected aligned performance outcomes from the approved competency curriculum. Engage an external academic reviewer to evaluate the curriculum and document the resulting instructional improvements. Surveys, disclosure records, and self-assessments may describe reach or practice, but they do not by themselves prove competency attainment. - *Owner*: Dean of the Graduate School. - *Artifact*: Published provision and reach figures, the aligned assessment method and selected performance results, the external reviewer's formal evaluation, and committee minutes documenting resulting curriculum revisions. - *Effort*: 6–9 months.

D4-people

Appendix A, read on this axis: at nascent, provision exists and is voluntary. At established, supervisors and examiners are trained. At leading, provision, reach, and selected aligned performance outcomes are reviewed by role and used for improvement.

D4-people → nascent — Appoint the academic lead with workload. Appoint a designated academic lead for AI research training competency and allocate formal time within the university workload allocation model. This is foundational work toward D4-people at nascent. The appointment alone does not advance the cell unless an institutional training offering is also available and the full A1 criterion is rescored Yes. Creating an unfunded coordination role typically results in a new webpage but no real instructional change. - *Owner*: Deputy Vice-Chancellor (Research), working with the Dean of the Graduate School. - *Artifact*: Formal role description, documented workload allocation in the faculty workload model, and official appointment date. - *Effort*: 6–10 weeks.

D4-people → established — Train supervisors and examiners, and make it a condition of appointment. Deliver structured training on the six core competencies to advisors and supervisors, alongside training on corresponding verification methods for examiners. Make completing this training a mandatory condition of examiner appointment, stated explicitly in official appointment letters. - *Owner*: Dean of the Graduate School for supervisors, and Chair, Graduate Studies Committee for examiners. - *Artifact*: Training completion registers by role, and updated examiner appointment letters containing the mandatory training condition. - *Effort*: 9–12 months.

D4-people → leading — Publish provision, reach, and aligned performance evidence by role. Publish evidence on provision, reach, and selected aligned performance outcomes by role on a regular schedule. The pack's graduate-student and supervisor surveys may describe reach and experience, but they do not by themselves demonstrate competency acquisition. Use the evidence in an externally informed curriculum-improvement cycle and document the resulting change. - *Owner*: Dean of the Graduate School. - *Artifact*: Published provision and participation figures by role, the aligned assessment method and selected performance results, external review evidence, and committee minutes recording the resulting program adjustment. - *Effort*: 3–4 months per survey cycle.

D4-systems

Appendix A, read on this axis: at nascent, individual completion is officially tracked. At established, approved delivery, aligned assessment, and retrievable records are integrated. At leading, coverage reports are generated automatically by cohort, department, and advisor population.

D4-systems → nascent — Track individual completion officially. Deliver AI-literacy instruction through a learning or administrative system that records completion for each student, advisor, and supervisor. Do not rely on an untracked open-access resource page as the evidence of completion. -

Owner: Director of Learning Technologies, working with the Graduate School. - *Artifact:* The tracked delivery record and an extract showing individual completions by role. - *Effort:* 6–10 weeks.

D4-systems → established — Support approved delivery and record completion. Deliver the approved competency curriculum through a supported institutional platform with versioned content, aligned assessments, and retrievable completion records for the relevant student and staff populations. A corpus-grounded tutor may support delivery, but it is optional rather than a required feature. - *Owner:* Director of Learning Technologies, working with the University Librarian. - *Artifact:* Supported delivery platform, approved curriculum version, assessment mapping, and retrievable completion records in institutional training databases. - *Effort:* 6–9 months.

D4-systems → leading — Automate the coverage reports. Configure institutional databases to generate recurring competency-coverage reports by student cohort, academic department, and advisor population without running a manual survey. - *Owner:* Director of Research Information Systems. - *Artifact:* Approved report definition, automated schedule, and the most recent output covering all three populations. - *Effort:* 4–6 months.

D4-process

Appendix A, read on this axis: at nascent, AI literacy appears in doctoral orientation and supervisor onboarding. At established, the curriculum is reviewed on a stated cadence. At leading, competency-outcome data is shared with sector benchmarking initiatives.

D4-process → nascent — Put AI literacy into orientation and onboarding. Introduce the six competencies in orientation for new doctoral students and in onboarding for newly appointed advisors and supervisors. - *Owner:* Dean of the Graduate School. - *Artifact:* Dated student-orientation agenda and advisor or supervisor onboarding checklist showing the AI-literacy component. - *Effort:* 4–8 weeks.

D4-process → established — Run the scheduled curriculum review. Review the AI methods curriculum at a stated cadence through a named committee, recording the tooling changes considered and the curriculum updates adopted. - *Owner:* Dean of the Graduate School, working with research methods coordinators. - *Artifact:* Approved review schedule, minutes from the most recent review, and the resulting curriculum changes. - *Effort:* 3 months for the first review, then recurring on cadence.

D4-process → leading — Contribute competency outcomes to sector benchmarking. Submit the institution's competency-outcome data to a named sector-level benchmarking initiative using its stated definitions and schedule. - *Owner:* Dean of the Graduate School, working with the institutional benchmarking lead. - *Artifact:* The named initiative, the submitted data specification, and the dated contribution record. - *Effort:* 3–6 months for the first contribution, then on the initiative's cadence.

Dimension 5 — Institutional benchmarking grid

D5-policy

Appendix A, read on this axis: at nascent, the cycle is articulated. At established, the institution scores itself against Dimensions 1–4 at a stated cadence and the scoring is reviewable by council, school and college boards, and the research-integrity committee. At leading, the scoring and its method are published.

D5-policy → nascent — Approve the benchmarking mandate. Draft a concise one-page mandate for the institutional AI-readiness benchmarking cycle (defining what is evaluated, who conducts it, how frequently it occurs, and which governance committee receives the reports) and submit it to [COMMITTEE] for formal approval. - *Owner:* Deputy Vice-Chancellor (Research). - *Artifact:*

Approved benchmarking mandate detailing evaluation scope, review cadence, and reporting lines, with the formal approval minute. - *Effort*: 3–6 weeks.

D5-policy → established — Adopt the grid as the stated method. Formally adopt the twenty-cell AI readiness grid in university policy as the official institutional assessment framework, define the documentary evidence standards for each cell, and complete an initial comprehensive assessment with attached evidence. - *Owner*: Deputy Vice-Chancellor (Research), working with the Chair of [COMMITTEE]. - *Artifact*: Policy clause naming the assessment framework and evidence standards, alongside the initial completed assessment report and its evidence register. - *Effort*: 3–5 months.

D5-policy → leading — Publish the scoring and the method. Publicly report the university's complete scoring alongside sufficient methodological detail for external reviewers to replicate the assessment, and incorporate this publication into regular institutional public reporting. The report notes that no university in its international sample currently publishes its self-assessment under this framework (§4.2), making the *leading* level a forward-looking standard that pioneering institutions can establish. - *Owner*: Deputy Vice-Chancellor (Research). - *Artifact*: Public website displaying all twenty cell ratings, evaluation methodology, supporting evidence, assessment date, and scheduled re-assessment date. - *Effort*: 2–3 months.

D5-people

Appendix A, read on this axis: at nascent, a named owner for the policy axis. At established, named owners for every axis of every dimension. At leading, external benchmarking or peer audit at a stated cadence.

D5-people → nascent — Appoint the owner of the cycle. Appoint a single accountable leader for the institutional benchmarking and improvement cycle, and publish their formal mandate. This owner may be a deputy vice-chancellor, a director of research integrity, or the chair of a standing committee. To supply evidence toward **D5-people → nascent**, the mandate must also assign accountability for the AI-in-research policy and its periodic review. It does not move Dimension 5 by itself. Dimension 5 reaches *nascent* only when all four D5 cells qualify at that level. - *Owner*: Vice-Chancellor makes the appointment, and the Deputy Vice-Chancellor (Research) or Director of Research Integrity holds the role. - *Artifact*: Formal appointment record and published mandate establishing the review cadence. - *Effort*: 2–4 weeks.

D5-people → established — Publish the twenty-cell ownership table. Assign a designated owning role to each of the twenty grid cells, and publish a comprehensive table listing the primary role, an appointed delegate, and a scheduled review date for every row. This requires an accountable assignment for every cell, which may be distributed among a smaller group of individuals, rather than only five dimension-level assignments. - *Owner*: Deputy Vice-Chancellor (Research). - *Artifact*: Published twenty-row ownership table listing role, delegate, and review date for each cell, with signed role acceptances on file. - *Effort*: 6–10 weeks.

D5-people → leading — Join or convene a peer benchmarking group. Establish a formal benchmarking consortium with [three to five] peer institutions to evaluate each other's evidence on a regular schedule under shared terms of reference, and publish the resulting peer review reports. - *Owner*: Deputy Vice-Chancellor (Research). - *Artifact*: Signed terms of reference for the peer benchmarking group, alongside the initial received peer evaluation report and management response. - *Effort*: 6–12 months.

D5-systems

Appendix A, read on this axis: at nascent, the axes are partially populated. At established, the records the four axes require sit in one retrievable place. At leading, the scoring is published in a form peers can compare.

D5-systems → nascent — Stand up the evidence register. Establish a single central repository that contains the documentary evidence supporting the current maturity level for each of the twenty cells. Index every artifact by cell, owning role, creation date, and review date so reviewers do not

have to gather the evidence piecemeal from different departments. - *Owner*: Director of Research Integrity. - *Artifact*: Central evidence repository and its completed twenty-row index, with document titles, owning roles, creation dates, and review schedules. - *Effort*: 4–8 weeks.

D5-systems → established — Consolidate the registers with retention and access rules.

Consolidate all related records (including impact assessments, the tool register, exception logs, training databases, and audit findings) into a single governed repository with defined data retention schedules and access controls. Then run and document a retrieval audit that requests the core artifacts, records what was produced, and measures the time required. - *Owner*: Chief Information Officer, working with the Data Protection Officer. - *Artifact*: Central repository index, documented retention schedule, formal data access policy, and the dated retrieval-audit record listing artifacts produced and elapsed retrieval time. - *Effort*: 4–6 months.

D5-systems → leading — Publish a machine-readable scoring extract.

Publish the university's maturity scoring as a structured data export (including cell identifier, level, evidence references, and timestamp) using an open schema with version tracking, published alongside the narrative report so peer institutions can conduct direct comparisons. - *Owner*: Director of Research Information Systems. - *Artifact*: Published open dataset with schema documentation and complete version history. - *Effort*: 2–3 months.

D5-process

Appendix A, read on this axis: at nascent, regulatory deadlines are recognised but not assigned. At established, a documented preparation plan exists for each applicable deadline and the scoring runs at cadence. At leading, execution results are published and the method is contributed to sector standardisation.

D5-process → nascent — Name the regulatory dates in the institutional plan.

Formally record in the institutional strategic plan every regulatory compliance date that applies to the institution's jurisdictions, whether or not responsibility has been assigned yet. This includes: (1) the EU AI Act Article 27 Fundamental Rights Impact Assessment obligation taking effect on August 2, 2026, where the institution is within Article 2's territorial scope and is a body governed by public law or a private entity providing public services that deploys a listed high-risk system, (2) the Australian Privacy Act 1988 automated-decision-making transparency obligation taking effect on December 10, 2026, for institutions subject to that regime, and (3) the local statutory equivalents in any other relevant jurisdiction. - *Owner*: General Counsel or University Secretary, working with the Data Protection Officer. - *Artifact*: Strategic plan entry detailing compliance dates, statutory obligations, and a formal institutional applicability determination. - *Effort*: 2–4 weeks.

D5-process → established — Run the cycle and the preparation plans at cadence.

Execute the self-assessment scoring on a regular schedule, backed by concrete preparation plans for each upcoming regulatory deadline (identifying designated owners, key milestones, and implementation status). Submit these materials to university council, school research committees, and the research integrity committee for formal review. - *Owner*: Deputy Vice-Chancellor (Research), reporting to [COMMITTEE]. - *Artifact*: Dated assessment report, preparation plans showing owners and milestone progress for each regulation, and minutes from all three committees confirming receipt. - *Effort*: 4–6 months to complete the first cycle.

D5-process → leading — Publish execution and contribute the method.

Publish an implementation review detailing actual institutional progress against each preparation plan after regulatory deadlines pass (including transparent explanations for any delayed milestones), and contribute the evaluation methodology to sector-wide standardization working groups. - *Owner*: Deputy Vice-Chancellor (Research). - *Artifact*: Published post-deadline evaluation report, alongside formal documentation of contributions submitted to sector standardization bodies. - *Effort*: 4–8 months.

Sequencing: 90 days and 12 months

The typical starting profile

The typical institution evaluated in the report's baseline sample publishes a broad AI policy touching on research integrity and reproducibility, but lacks any codified competency expectations, has no examiner-side rules, and conducts no regular benchmarking. On the maturity grid, that typical profile appears as follows:

Dimension	Level	Binding axis	Why
D1 — Human-in-the-loop discipline	nascent	systems, process	The position is stated, but no form enforces it.
D2 — Responsible use in practice	nascent	people, process	A disclosure template exists, but nothing verifies compliance.
D3 — Tooling that promotes responsible use	nascent	systems	The tool register exists, but protected enterprise deployments and published session-log retention are incomplete.
D4 — AI-literate humans	absent	policy, people, systems, process	Training provision remains voluntary and unintegrated.
D5 — Institutional benchmarking grid	absent	all four	There is no designated owner, regular review cycle, or evidence register.

Table 7. Baseline maturity profile showing dimension levels, binding axes, and explanatory rationales.

Binding dimension overall: Dimension 4, which is the lowest-numbered dimension sitting at the minimum maturity level (absent). **Assurance status:** Because Dimension 5 is below the established level, no maturity claim above nascent on any dimension is assured.

This profile serves purely as an illustrative example rather than an institutional default. You should calculate your university's specific baseline using the triangulation table in §1.7 before applying the sequencing schedules below, and skip any rungs that your assessment shows are already completed.

The natural instinct for many leaders in this situation is to start immediately with Dimension 4, simply because it is the weakest area. However, in my experience, that is usually the wrong first move. Achieving the *established* level on Dimension 4 requires long academic-year governance cycles (such as formal curriculum approval taking 6 to 9 months, or comprehensive supervisor and examiner training taking 9 to 12 months). Furthermore, none of those educational achievements can be formally evidenced without the Dimension 5 register, which does not exist yet. I strongly recommend starting with the faster, low-cost rungs that build the evidence and governance foundation that all other dimensions will report into.

Days 1–90

Seven priority work packages toward *nascent* rungs can be completed comfortably within a single quarter. Together, they require several weeks of dedicated professional staff time but no capital expenditure. Their completion triggers a rescore rather than an automatic level change.

Window	Work package	Owner	Artifact due
Days 1–30	D5–people → nascent	Vice-Chancellor appoints	Appointment record and published mandate
Days 1–30	D5–process → nascent	General Counsel with DPO	Plan entry naming each regulatory date and whether it applies

Window	Work package	Owner	Artifact due
Days 1–30	D3-policy → nascent	Chief Information Security Officer	Approved interim data rule and register of qualifying deployments
Days 15–45	D5-policy → nascent	Deputy Vice-Chancellor (Research)	Approved benchmarking mandate
Days 20–90	D3-systems → nascent	Chief Information Officer	Published register of AI tools in research use
Days 30–75	D1-policy → nascent	Deputy Vice-Chancellor (Research)	Approved one-page position statement
Days 45–90	D5-systems → nascent	Director of Research Integrity	Evidence register, twenty rows populated

Table 8. Implementation schedule for days 1 to 90 specifying work packages, owners, and required artifacts.

By day 90, the institution has established an accountable owner, an approved benchmarking mandate, a complete register of AI tools currently in use, a clear rule protecting unpublished research data, a published institutional position, and a central evidence register. These work packages support a D5 nascent rescore. Dimension 5 moves from *absent* to *nascent* only if every applicable A1 indicator at the nascent rung across all four D5 cells is then answered Yes.

Months 4–12

Window	Work packages	Primary owner	What it buys
Months 4–6	D1-people → nascent, D2-policy → nascent, D4-policy → nascent, D1-systems → nascent, D1-process → nascent	Graduate School and Research Integrity	Supporting work across five cells before the next A1 rescore
Months 4–9	D3-policy → established, D3-people → established	Chief Information Officer	A published procurement standard and a panel with authority to refuse
Months 5–10	D1-policy → established, D2-policy → established	Research Integrity Committee	The labor-versus-judgment schedule and the four mode rules, approved together
Months 6–12	D3-process → established	Data Protection Officer	The impact assessments the institution's jurisdictions require, in one register
Months 6–12	D4-policy → established, toward D4-people → nascent, D4-systems → nascent	Dean of the Graduate School	The six competencies approved through the curriculum, with foundational D4-people work that remains incomplete until a training offering is present
Months 9–12	D5-people → established, D5-policy → established, D5-process → established	Deputy Vice-Chancellor (Research)	An accountable assignment for every cell, the grid adopted as the stated method, the first full scored cycle with preparation plans

Table 9. Implementation plan for months 4 to 12 specifying work packages, owners, and expected outcomes.

At the end of twelve months, the listed work packages are expected to provide evidence across the policy axis, assign owning roles, and populate the evidence register. They do not predetermine any dimension's later state. Record each later state only after a complete A1 rescore. Under this schedule, D4-process remains unaddressed, and D4-people remains conditional on an institutional training offering being available, so Dimension 4 cannot advance on the schedule alone. The Dimension 5 assurance status must likewise be determined from the complete rescore of all four D5 cells.

Year two then focuses on completing the people, systems, and process rungs needed for the *established* transition. You should view this twelve-month plan as building the foundational evidence base that those subsequent rungs will rely on, rather than expecting to reach *established* across entire dimensions immediately. Any plan that claims to reach established overall within twelve months is almost certainly averaging across its axes rather than applying the minimum rule.

If your profile differs

Policy-heavy institutions (where the policy axis is at established, but other axes remain nascent). You should skip additional policy drafting entirely. Spend the first quarter focusing on D1-systems → established, D2-process → nascent, and D5-systems → nascent. The primary operational bottleneck here is that while the policy exists, no formal administrative system or workflow requires anyone to follow it.

Practice-heavy institutions (where the people axis is well ahead of formal policy). Focus on codifying existing practices rather than designing new procedures from scratch. Advance through D1-policy → nascent, D2-policy → nascent, and D4-policy → nascent by documenting the good work already being done on the ground and submitting it to [COMMITTEE] for formal approval. This generally takes only six to ten weeks, and it produces a far more practical and grounded policy than a committee starting from a blank page.

Regulator-pressed institutions (facing binding statutory compliance deadlines within twelve months). Sequence D5-process → nascent, D3-policy → nascent, D3-systems → nascent, D3-process → nascent, and D3-process → established first, in that exact order. The cumulative progression rule does not bend for a looming deadline. A formal impact assessment cannot be properly scoped until you have an accurate register of the AI tools being used, and building that tool register is a six-to-ten-week project that institutions often discover too late.

Five ways this analysis goes wrong

1. **Averaging across the axes.** A dimension's maturity level is strictly the minimum score across its four axes. Averaging them together converts an absent axis into an artificial, comfortable middle score, which hides the exact operational bottleneck the grid was designed to locate.
2. **Generating a single institutional score.** There is no single, overall institution-wide score, and you should never calculate one. A single summary number will inevitably be quoted in meetings for years without providing anyone with actionable guidance on what needs fixing.
3. **Treating every A2 and A3 item as the same kind of evidence.** A report of whether a milestone conversation occurred, a knowledge probe, a training-exposure item, and a belief item answer different questions. Keep them separate, report their coverage, and use only a directly aligned operational item to test an A1 criterion.
4. **Attempting to climb two rungs at once.** Every established rung is designed assuming that the corresponding nascent artifact is already in place. Trying to skip ahead usually produces an approved policy document with no operational systems underneath it, leaving the cell in the exact same position during the next review cycle.

5. Purchasing AI software tools before establishing the policy rules. Dimension 3 evaluates tools against operational definitions of responsible use that must come from Dimensions 1 and 2. Procuring tools before defining those standards means relying entirely on vendor claims, which is precisely why institutions get stuck at the *nascent* level on D3-policy.

Crosswalk: Instrument item to report section

Item	Report section
§1.1 The three sightlines	Part 4 §4.1, Appendix A
§1.2 Step 2: The cumulative progression rule (the ratchet)	Appendix A ("All of <i>established</i> , plus...")
§1.2 Step 3: Criterion-specific survey evidence at established and above	Part 4 §4.1 (definitions of <i>nascent</i> and <i>established</i>)
§1.2 Step 5: Undocumented practice records at nascent	Part 4 §4.1 ("operating at a reviewable cadence")
§1.3 Binding axis (minimum across axes)	Part 2 §2.5, Part 4 §4.1
§1.4 Binding dimension (dimension sequence)	Part 2 preamble, Part 4 §4.3
§1.4 The Dimension 5 interlock	Part 4 §4.1
§1.5 No single institution-wide score	Part 4 §4.3
§1.5 Class A–D placement is separate	Part 1 §1.3
§1.6 Worked interpretation patterns A–D	Part 4 §4.3 (worked institutional examples)
§1.7 Triangulation table	Part 4 §4.1, Appendix A
§1.8 Dimensional profile output	Part 4 §4.3
Rungs D1-policy, D1-people, D1-systems, D1-process	Part 2 §2.1, Appendix A row 1, Appendix G, Part 5 §5.3
Rungs D2-policy, D2-people, D2-systems, D2-process	Part 2 §2.2, Appendix A row 2, Part 5 §5.2
Rungs D3-policy, D3-people, D3-systems, D3-process	Part 2 §2.3, Appendix A row 3, Appendix C, Part 5 §5.1, §5.4
Rungs D4-policy, D4-people, D4-systems, D4-process	Part 2 §2.4, Appendix A row 4, Part 5 §5.2, §5.3
Rungs D5-policy, D5-people, D5-systems, D5-process	Part 2 §2.5, Appendix A row 5, Part 5 §5.1
The four AI-use modes in D2-policy → established	Part 2 §2.2
The six tool properties in D3-policy → established	Part 2 §2.3
The six competencies in D4-policy → nascent and → established	Part 2 §2.4
The labor-versus-judgment schedule in D1-policy → established	Part 2 §2.1, Appendix G
Regulatory dates in D3-process → established and D5-process → nascent	Part 2 §2.3, Part 2 §2.5, Part 4 §4.2
Resolvability testing in D3-systems → leading	Part 2 §2.3, Appendix C Classes 1 and 6
Defend-your-use rule in D1-process → established	Part 5 §5.3, Appendix B (Canada, University of Toronto formulation)

Item	Report section
90-day and 12-month sequencing	Part 5 §5.1–§5.4
Five ways this analysis goes wrong	Part 4 §4.1, Part 4 §4.3

Table 10. Mapping of assessment instrument items to their corresponding report sections.

instats

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

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

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

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.