Assessment Guidebook – A practical guide for innovative and future-oriented assessment in the era of AI

XJTLU | EDU

Created by EDU Assessment Working Group:

Ling Wang, Joseph Tinsley, Wenhao Wang, Nauman Saeed

Version 1.0

Table of Contents

Acknowledgement 1

Section A: Background 2

Section B: Guiding principles for assessment in the era of AI 2

Section C Tripartite Assessment Model Framework 5

C1. Tripartite Assessment Model for the AI Era (TAMA) 6

A Human Intelligence (The Uniquely Human) 6

B Artificial Intelligence (Human Use of AI) 7

C Synergistic Intelligence (Human AI Integration) 8

C2. Implementation framework 9

Section D: Assessment Recipes 9

D1. Large-classes assessment recipes 9

Recipe 1 Swarm Intelligence Assessment 9

Recipe 2 Invisible Assessment Ecosystem 11

Recipe 3 Peer-reviewed Digital Portfolios 12

Recipe 4 Adaptive quizzes with tiered challenges 13

Recipe 5 Dynamic Concept Mapping 13

Recipe 6 AI-supported tiered assessment architecture 14

Recipe 7 Peer + AI feedback ecosystem 15

Recipe 8 Group project + individual oral defence with independent contribution multiplier 15

D2. Project-based Learning (PBL) assessment recipes 17

Recipe 1 Competency and milestone-based project assessment 17

Recipe 2 The project that never ends 18

Recipe 3 Client-focused project deliverables 20

Recipe 4 PBL via collaborative platforms 21

Recipe 5 Multi-agent AI project studio 21

Recipe 6 Challenge-based community innovation project 22
Recipe 7 Multi-staged assessment model 23
Recipe 8 AI-orchestrated learning intelligence framework 25
Section E: Assessment redesign implementation plan 28
Phase 1: Foundation & Capacity Building (September 2026-January 2027) 28
Phase 2: Piloting & Evidence Generation (January 2027-August 2027) 28
Phase 3: Scaling & Integration (September 2027-January 2028) 29
References 31

Acknowledgement

This guidebook has been developed through the collective efforts of colleagues across the University who are committed to reimagining the future of assessment in the era of artificial intelligence. We would like to express our sincere appreciation to the University's Assessment Champions, including module leaders, programme directors, and teaching staff from a wide range of disciplines, who generously contributed their ideas, insights, and perspectives throughout the development of this guidebook.

While informed by current developments in assessment, educational innovation, and AI, this guidebook is not intended to be a collection of existing practices. Instead, it seeks to explore new possibilities and stimulate thinking about what assessment could become in an increasingly AI-enabled future. Many of the concepts, models, and assessment recipes presented in this guide emerged from discussions that challenged conventional assumptions about assessment and encouraged more imaginative, ambitious, and future-oriented approaches to student learning and evaluation.

We gratefully acknowledge the contributions of the following University Assessment Champions (listed alphabetically by surname)

Khalid Akhal (Entrepreneurship), Bryan Jia Hao Chai (Strategic Management and Organisations), Biwei Cong (Film and Creative Technology), Juan Carlos Dall'Asta (Architecture), Yue Li (Computing), Na Li (Future Education), Marco Pellitteri (Media and Communication), Michael Prieler (Media and Communication), Stefan Seedorf (Entrepreneurship), Thomas Selig (Computing), Ting Ting Tay (Future Education), Rong Wang (Future Education), Quan Zhang (Mechatronics and Robotics), Peng Zhao (Environmental Science), Lu Zong (Maths and Physics)

Section A: Background

In response to the rapid emergence of GenAI and its transformative implications for higher education, the University has positioned Education + AI as a major institutional strategic priority. The university's Education + AI Strategic Framework highlights the need to redesign curriculum, pedagogy, and assessment to prepare graduates for an increasingly AI-mediated world.

Within this context, assessment innovation has become a strategic priority in the transformation of education. XJTLU has a clear institutional intention to significantly reduce reliance on traditional examinations and conventional academic reports across disciplines wherever pedagogically appropriate. This direction is not intended to eliminate rigour, academic standards or disciplinary knowledge. Instead, it reflects a broader commitment to redesigning assessment so that it better supports meaningful learning, graduate capability development, creativity, interdisciplinarity, and real-world problem solving in the era of AI. Particular emphasis has been placed on two educational contexts that are increasingly central to contemporary higher education: Project-Based Learning (PBL) and large-class teaching. PBL reflects the university's wider commitment to interdisciplinary, experiential, and industry-engaged learning, while large-class teaching reflects the practical realities of delivering high-quality education at scale across diverse disciplinary areas. Both contexts require assessment approaches that are educationally meaningful, scalable, inclusive, sustainable for staff workload, and robust within an AI-enabled learning environment.

This Assessment Guidebook has therefore been developed as a practical institutional guide, aiming to support module leaders, programme directors, and teaching teams to develop innovative assessment approaches that align with the XJTLU's Education + AI Strategic Framework while remaining academically rigorous, discipline-sensitive, and operationally realistic.

Section B: Guiding principles for assessment in the era of AI

Principle 1 - Assessment for learning

What this principle means

Assessment should not be viewed simply as a way of measuring what students have learned at the end of a module. In the AI era, assessment needs to play a much stronger role in supporting learning throughout the learning journey.

Students learn most effectively when they receive opportunities to practise, receive feedback, reflect on their performance, and improve their work. Assessment should

therefore be designed not only to evaluate achievement, but also to help students develop knowledge, skills, and confidence over time. This principle encourages academics to think of assessment as an integral part of learning rather than a separate activity that takes place after learning has occurred.

Implications for assessment design

Assessment should: incorporate iterative feedback loops; enable revision and resubmission; reward improvement and development over time; embed reflection and self-evaluation; provide opportunities for students to learn from failure and experimentation.

Principle 2 – AI as cognitive partner

What this principle means

Gen AI is becoming a normal part of study, work, and professional practice. Rather than designing assessment on the assumption that students should never use AI, assessment should help students learn how to use AI effectively, critically, and responsibly.

Students will increasingly work alongside GenAI systems in their future careers. They need opportunities to learn when GenAI can be helpful, when it should be questioned, and how its outputs should be evaluated and improved. The goal is not to replace human thinking with AI, but to help students develop the ability to work productively with AI while maintaining responsibility for decisions, judgement, and outcomes.

Implications for assessment design

Assessment should: explicitly incorporate AI where appropriate; evaluate students' interactions with AI systems; assess prompt design, critique, verification, and refinement; require students to justify AI-assisted decisions; focus on evaluative judgement and critical reasoning rather than production alone.

Principle 3 – Authenticity

What this principle means

Students are more engaged when assessment reflects the kinds of problems, situations, and challenges they are likely to encounter beyond university. In the AI era, authenticity becomes even more important. Tasks that simply ask students to reproduce information or generate generic content can often be completed easily with AI tools. Assessment should therefore focus on meaningful disciplinary application, professional judgement, and engagement with realistic contexts. Authentic assessment helps students understand why their learning matters and how it connects to future careers, professional practice, and societal challenges.

Implications for assessment design

Assessment should: use real-world or realistic scenarios; engage external stakeholders where possible; incorporate AI-rich professional environments; evaluate decision-making under uncertainty; require disciplinary application rather than content reproduction.

Principle 4 – Process-focused

What this principle means

Traditional assessment often focuses primarily on the final product, such as an essay, report, presentation, or examination answer. However, in an AI-enabled world, the process through which students arrive at a solution would be as important as the final outcome.

Students should be encouraged to demonstrate how they approached a problem, what decisions they made, how they responded to feedback, how they used AI tools, and how their thinking evolved over time. This principle recognises that learning is often visible in the journey rather than only in the destination.

Implications for assessment design

Assessment should: capture learning journeys and development; require documentation of decision-making processes; include drafts, iterations, and developmental evidence; reward reflection, adaptation, and improvement; assess how students work, not only what they produce.

Principle 5 – Development of human capabilities

What this principle means

As AI becomes increasingly capable of generating text, solving problems, analysing information, and automating routine tasks, universities must place greater emphasis on developing the capabilities that remain uniquely human. These include: critical thinking; ethical judgement; creativity; collaboration; communication; empathy; leadership; adaptability; and self-awareness.

Assessment should help students develop and demonstrate these capabilities rather than simply rewarding information recall or routine content production. The question is no longer only: "What can students know?", but increasingly: "What can students do that adds value beyond what AI can do?"

Implications for assessment design

Assessment should: prioritise judgement over information recall; reward creativity and innovation; evaluate collaboration and communication; assess ethical and societal

reasoning; promote self-awareness and metacognition; focus on human strengths within human-AI partnerships.

Principle 6 – Future-oriented

What this principle means

Today's students will graduate into a world that is changing rapidly. Many of the technologies, jobs, and challenges they will encounter may not yet exist. Assessment should therefore prepare students not only for current professional practice but also for future uncertainty. Students need opportunities to explore emerging technologies, anticipate future challenges, think creatively, and develop the adaptability required to thrive in changing environments. Future-oriented assessment encourages students to become active contributors to shaping the future rather than simply responding to it.

Implications for assessment design

Assessment should: address emerging societal and technological challenges; encourage speculative and futures thinking; incorporate uncertainty and ambiguity; assess adaptability and resilience; promote innovation and responsible experimentation; evaluate students' overall capacity to shape future human-AI ecosystems.

Section C Tripartite Assessment Model Framework

The TAMA (Tripartite Assessment Model for the AI-Era) Framework is a reimagining of how XJTLU evaluates learning in an era where artificial intelligence has become an integral part of students' educational journey. In line with the Education + AI strategy, AI is understood as a means to enhance student's demonstration of learning in assessments, rather than a threat to academic integrity or assessment validity. As a result, TAMA recognises that assessment must account for and evaluate sophisticated human-AI collaboration (Xi, 2025).

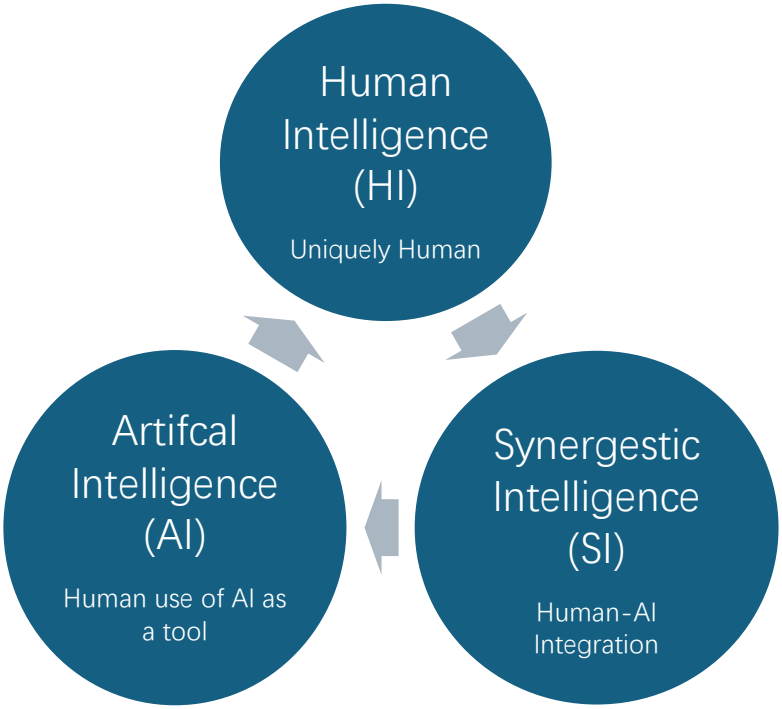
To do so, the framework establishes three interconnected dimensions of assessment, all of which involve engagement with AI:

Human Intelligence (HI): Assessing the uniquely human capabilities that remain essential for students in education and employment

Artificial Intelligence (AI): Assessing technical proficiency and critical engagement with AI tools within specific domains

Synergistic Intelligence (SI): Assessing the emergent value created through human-AI integration

C1. Tripartite Assessment Model for the AI Era (TAMA)



A Human Intelligence (The Uniquely Human)

Human intelligence capabilities are those which are uniquely human and involve the demonstration of ethical, aesthetic or emotional judgement. Assessment tasks may involve interaction with AI, but the final evaluative judgements are made by students.

A1 Assessment Competencies

HI 1: Critical Judgement	• Ethical reasoning • Contextual adaptation • Value-based decision making
HI 2: Creative Synthesis	• Original idea generation • Cross-domain connections • Metaphorical thinking
HI 3: Emotional Intelligence	• Empathy and perspective-taking • Interpersonal dynamics • Cultural sensitivity
HI 4 Metacognitive Awareness	• Self-regulation • Learning strategy adaptation • Critical reflection

A2 Assessment methods

The assessment methods for human intelligence capabilities require scenarios in which students are able to demonstrate their ethical, creative and metacognitive judgement. Examples may include

- Ethical dilemma analyses with AI-generated scenarios
- Creative remix tasks combining AI outputs in novel ways
- Metacognitive reflections on AI usage
- Live simulation scenarios where AI provides data but human judgment makes decisions

B Artificial Intelligence (Human Use of AI)

Artificial intelligence competencies are competencies which involve the interaction between students and artificial intelligence. These competencies are a combination of discipline specific technical prophecies and broader AI literacy competencies. Assessment tasks involve students demonstrating technical proficiency and the ability to use AI critically and responsibly.

B1 Assessment Competencies

AI 1: AI Literacies	• Understanding AI capabilities • Recognizing biases in training data • Articulating appropriate use cases
AI 2: Technical Proficiency	• Prompt engineering • Tool selection and combination • Output interpretation
AI 3: Critical Evaluation	• Assessing output quality • Identifying hallucinations • Validating AI-generated information
AI 4 Ethical Application	• Responsible use of AI • Transparency in using AI • Citation and acknowledgment of AI

B2 Assessment methods

The assessment methods for artificial intelligence capabilities require scenarios in which students are able to demonstrate their critical and responsible engagement with AI tools and process. Examples may include

- AI audit assignments in which students analyze and critique AI outputs
- Prompt optimization tasks where students iteratively improve AI interactions
- Tool selection justifications in which students defend their choice of AI tools for specific tasks

C Synergistic Intelligence (Human AI Integration)

Artificial intelligence competencies are competencies which involve the synthesis of student intelligence and artificial intelligence in more sophisticated, long-term processes. These are strategic and transformative competencies which students demonstrate through project and portfolio type assessment tasks.

SI 1: Integration Capacity	• Strategic tool deployment • Workflow design • Timing of AI intervention
SI 2: Critical Synthesis	• Value-added transformation • Integration of multiple sources • Enhancement of AI outputs
SI 3: Meta-Collaboration	• Self-awareness in partnership • Adaptive collaboration strategies • Ethical positioning in delegation

Assessment methods

The assessment methods for synergistic intelligence capabilities require scenarios in which students are able to demonstrate their interaction with AI over longer time period. Examples may include

- Evolving cases studies where students work with a case study that changes based on their AI interactions
- Collaboration process mapping in which students document decision points in human-AI workflow
- Value-added analysis in which students demonstrate how AI has enhanced human thinking

C2. Implementation framework

	Assessment Methods	Evaluation Methods
Human (HI)	Reflective essays Ethical case analyses Live discussions	Depth of reasoning Value articulation
Artificial (AI)	Prompt engineering logs AI output critiques Tool selection justification	Technical proficiency Critical evaluation
Synergistic (SI)	Process portfolio Workflow design Collaboration Maps	Value creation Integration quality

These assessment methods are presented as broad categories for the purposes of this document. In practice, assessments will combine multiple methods to best assess the three competency types. EDU will support colleagues in developing assessment tasks and new rubrics and marking descriptors to assist colleagues in evaluating student competencies.

Section D: Assessment Recipes

D1. Large-classes assessment recipes

Recipe 1 Swarm Intelligence Assessment

- **Purpose**

It is designed for a large student cohort collectively operating as a distributed intelligence network attempting to solve a highly complex huge evolving challenge that cannot be solved effectively by individuals, small groups or AI systems alone. It shifts from traditional one student by one script with one mark, to network-based, non-linear, ecosystmemic. The cohort of students becomes the 'intelligent organism' in the assessment ecosystem, with each student acting as one node within the larger adaptive system.

It aims to move beyond traditional assessment approaches that focus primarily on individual knowledge reproduction, isolated problem-solving, or static assignment submission. Instead, it aims to evaluate students' ability to operate within large-scale human-AI collaborative ecosystems where intelligence is distributed, dynamic, adaptive, interdisciplinary, and continuously evolving.

It tries to assess students' ability to: contribute meaningfully to collective intelligence systems; collaborate across large distributed networks; govern and interact with AI-generated systems and recommendations; make decisions under uncertainty; negotiate competing priorities and stakeholder interests; exercise ethical, strategic and system-level judgement; and operate effectively within future human-AI professional environment.

- **Class Size**

50 - 400 students

- **Suitable Disciplines**

Science, pharmacy, design, maths and physics, computer science, engineering and artificial intelligence, business economics and management, film and creative arts, humanities and social sciences, education.

- **Assessment Examples**

Typical assessment formats may include:

Live simulation dashboards - students to monitor evolving systems through dynamic interfaces;

Distributed governance roles - each student controls one part of the larger ecosystem;

AI-generated scenarios and ongoing disruptions - AI continuously introduce new challenges, failures or conflicts, disruptions, ethical dilemmas;

Intervention decision rounds - student make strategic decisions which may affect the entire huge system

Cross-collaboration tasks – students negotiate and exchange information across teams

Stakeholder negotiation simulations – students represent competing interests or priorities

Collective (re)design challenges – students collaboratively redesign failing systems

Human override decisions – students determine when AI systems should be challenges, ignored, or overridden.

Reflective governance analysis – students explain how human judgement has (re)shaped system outcomes.

Module leaders can determine who can communicate with whom; which information is visible or hidden; which systems depend on each other; and where uncertainty exists. Students can be distributed across functional roles, such as: data analysts, AI governance officers, infrastructure coordinators, ethics councils, public communication teams, financial

regulators. They do not need to submit assignments individually. Instead, they need to monitor systems; interpret AI-generated outputs; make intervention decisions; negotiate with other network nodes; exchange intelligence; respond to crises; improve or redesign systems; justify governance decisions; adapt to dynamic disruptions.

Students will be assessed on: quality of judgement; systems thinking; collaborative intelligence; adaptability; governance; communication; ethical reasoning; strategic intervention; contribution to collective problem-solving.

Recipe 2 Invisible Assessment Ecosystem

- **Purpose**

Assessment is continuously embedded within learning activity rather than separated into visible assessment events such as exams, assignments, quizzes or reports. Instead of repeatedly asking students to “submit work of assessment”, this recipe is designed as a live intellectual activity ecosystem in which evidence of learning emerges naturally through: reasoning, modelling, interpretation, collaboration, explanation, decision-making, critique, adaptation and interaction with peers and AI.

It aims to assess: how students think rather than simply what they produce; how students operate within AI-rich disciplinary environments; how students interpret, challenge, refine, and govern AI-generated outputs; how students reason under uncertainty and changing conditions; how students communicate and collaborate intellectually; how students continuously develop understanding over time. It shifts from evaluating isolated products, towards observing intellectual behaviour in action across time.

- **Class Size**

50 - 300 students

- **Suitable Disciplines**

All disciplines

- **Assessment Examples**

Maths & Physics as an example

Module leader designs scientific experiences; modelling environments; interaction structures; AI behaviours; intellectual tensions; collaborative dynamics. Weekly scientific engagement environments are organized which can be interactive scientific studios, equation interpretation arena, computational activity ecosystems, mathematical thinking arenas that allow students to continuously participate.

For equation interpretation arena, students work in teams either in a large active-learning classroom OR online collaborative platform. Each table has shared digital workspace, AI modelling assistant, collaborative whiteboard, and equation visualization tools.

Students act as wave-system analyst inside a future quantum communications lab. They are provided with some AI-produced wave equations and AI claims all the models accurately predict the wave system. They need to test assumptions, visualize equations, challenge and debate each other, ask AI for simulations, compare physical realism, perform scientific reasoning, and undertake peer scientific communication.

Evidence for assessment may include: scientific decision logs by each team about which equation they rejected and why, and how their thinking changed; whiteboard capture derivations, annotations, corrections, and model revisions; AI interaction history.

Recipe 3 Peer-reviewed Digital Portfolios

- **Purpose**

To provide personalized, scalable feedback in large classes while developing students' evaluative skills. This redesign addresses the challenge of giving meaningful feedback to large cohorts by leveraging peer learning and structured evaluation frameworks.

- **Class Size**

100+ students. The system becomes more effective with larger cohorts as it creates a richer pool of peer reviewers.

- **Suitable Disciplines**

Humanities, Social Sciences, Education, Business, Communications, Design, and any field where process and development can be documented.

- **Assessment Examples**

Weekly Reflection Portfolios: Students submit 3-4 key learnings each week. Peers provide feedback using a rubric. Instructor provides feedback on common strengths and areas for improvement.

Project Development Logs: For a semester-long project, students document their process bi-weekly. They review 2 peers' logs at mid-term and final, focusing on progress and methodology.

Digital Skill Portfolios: Students collect evidence of skill development (e.g., data analysis, writing samples). Peers evaluate against program-level competency rubrics.

Recipe 4 Adaptive quizzes with tiered challenges

- **Purpose**

To replace traditional high-stakes exams with personalized assessment pathways that adjust to student performance levels, reducing academic anxiety while maintaining rigor in large cohorts.

- **Class Size**

150+ students. Technology-enabled scaling makes this feasible where manual differentiation would be impossible.

- **Suitable Disciplines**

STEM fields, Quantitative Social Sciences, Languages, Health Sciences, Engineering.

- **Assessment Examples**

Branching Case Quizzes: Using the VLE, initial case questions are determines pathway. Correct answers unlock advanced application questions; incorrect answers lead to foundational concept questions. All students attempt same number of questions but at appropriate difficulty.

Competency-Based Checkpoints: Students must demonstrate mastery of "Level 1" concepts before accessing "Level 2" application questions. Automated system allows unlimited attempts at each level with different question variants.

Recipe 5 Dynamic Concept Mapping

- **Purpose**

To assess deep conceptual understanding and systems thinking through visual knowledge representation. This approach moves beyond assessing students recall abilities to evaluate how students identify conceptual relationships and build connections between different knowledge bases.

- **Class Size**

150-300 students. This size is large enough to generate diverse mapping approaches for comparison and peer learning, while remaining manageable with digital tools and sampling strategies.

- **Suitable Disciplines**

Biology, Chemistry, Physics, Engineering, Environmental Science, Computer Science and any

discipline with complex interconnected concepts.

- **Assessment Examples**

Assessment Mapping Portfolio: Students use digital mind mapping tools (e.g. Miro or Canva) to create a portfolio of mind maps

• Map 1 (Week 4): Foundational module concepts

• Map 2 (Week 8): Mid-level integration drawing on knowledge from across the programme

• Map 3 (Week 12): Complex synthesis involving theoretical and practical elements

Recipe 6 AI-supported tiered assessment architecture

- **Purpose**

To redesign assessment for large cohorts by distributing assessment into low-stakes, medium-stakes, and high-value authentic tasks, reducing marking overload while maintaining deep learning and meaningful engagement.

- **Class Size**

100 – 500 students

- **Suitable Disciplines**

Engineering, Business, Computer and engineering, education, foundation programmes

- **Assessment Examples**

1. Weekly AI-assisted concept checks + final authentic case study

Students complete short weekly AI-supported quizzes or scenario analyses that provide immediate formative feedback. These combine together to a larger authentic task where students apply accumulated knowledge to solve a real-world problem. The focus is on demonstrating synthesis rather than memorisation.

2. Automated coding quizzes + group industry project

Students undertake auto-graded coding or technical exercises throughout the semester to build foundational skills. At the same time, they collaborate on an industry-style project requiring design decisions, justification of approaches, and critical evaluation of AI-generated solutions.

3. AI-generated formative feedback tasks + oral defense viva

Students submit iterative drafts that receive AI-supported feedback before final submission.

To ensure authenticity and deeper understanding, students participate in a short oral defence where they explain their reasoning, design choices, and use of AI tools.

Recipe 7 Peer + AI feedback ecosystem

- **Purpose**

To create sustainable feedback systems in large classes by combining structured peer review with AI-generated formative feedback and tutor moderation.

- **Class Size**

100 – 500 students

- **Suitable Disciplines**

All disciplines, particularly effective in writing-intensive or studio-based programmes

- **Assessment Examples**

1. Peer-reviewed essays with AI feedback comparison

Students first receive AI-generated feedback on draft essays before participating in structured peer review activities. They then compare the strengths and limitations of peer versus AI feedback and explain how both informed their revisions.

2. Studio critiques supported by AI-generated suggestions

Students upload visual or design work to receive AI-generated critique relating to structure or user experience. These suggestions are then discussed in peer critique sessions where students evaluate which feedback is valuable, superficial, or misleading.

3. Business pitch reviews using peer + AI evaluation matrices

Student teams present business proposals evaluated by peers and AI-supported scoring frameworks aligned to assessment rubrics. Reflection tasks require students to analyse discrepancies between human and AI evaluations and identify effective human-centred communication.

Recipe 8 Group project + individual oral defence with independent contribution multiplier

- **Purpose**

To enhance authenticity in group project assessment and ensure individual accountability, a small percentage of individual oral defense component is added. The multiplier adjusts the group project score based on each student's demonstrated independent contribution, making the final grade more accurately reflect individual performance.

- **Class Size**

100-250 students

- **Suitable Disciplines**

Maths & physics, Engineering, Science

Assessment Examples

Financial maths as example

Phase 1 – Group Project

Students form groups of 4–5 to conduct a project on factor investing. The project assesses: Understanding of asset pricing theory; Coding skills; Academic report writing

Phase 2 – Individual Oral Defense

Upon submission of the group report, an oral defense is organised. Each student is assessed individually by two examiners on: Their understanding of the module's learning outcomes; Reflection on the project. Each student answers 4–5 questions of progressive difficulty.

Phase 3 – Mark Adjustment Based on Defense Performance

The independent contribution multiplier adjusts the group report mark as follows:

Oral Defence Score. Adjustment to Report Mark:

Below 40 (very poor) –30%; 40–49 (poor) –15%; 50–74 (median to good) No change; 75–85 (excellent) +5%; Above 85 (very excellent) +10%

Safeguards:

- If the multiplier would cause the final module mark to fall below 40 solely due to the adjustment, the mark is raised to 40 (minimum passing threshold).

- If the multiplier would cause the final module mark to exceed 100, the mark is capped at 100.

These safeguards ensure fairness while preserving the multiplier's purpose of evaluating individual contribution and understanding.

D2. Project-based Learning (PBL) assessment recipes

Recipe 1 Competency and milestone-based project assessment

- **Purpose**

To replace time-based progression with venture-maturity progression, aligning assessment with authentic entrepreneurial and innovation ecosystems. It encourages iterative development rather than one-off submission, and rewards experimentation, adaptation and evidence-based decision making. It aims to prepare students for uncertain, rapidly evolving future economies.

- **Class Size**

50-300 students. 300+ students will require strong venture clusters and mentor networks

- **Suitable Disciplines**

Entrepreneurship; Business, finance, economics and management, Artificial intelligence, Pharmacy, Design.

- **Assessment Examples**

Entrepreneurship as example

Milestone 1 discovery and opportunity validation –

Students do not need to submit a business plan. Instead, they must produce evidence that real people experience the problem, existing solutions fail, and the problem is worth solving. Evidence to provide: assumptions that proved wrong, misconceptions they held, and evidence that destroyed their first concept.

Milestone 2 strategic intelligence and market validation –

Students do not need to present slides. Instead, they will take the interrogation from a panel of users, customers, investors, industry partners, and AI market agents who will challenge students' assumptions, value proposition, market logic, and customer understanding. Before the interrogation, students will conduct market validation; perform competitor intelligence; analyse future trends; test customer desirability; and refine value propositions.

Evidence to provide: customer validation evidence; market intelligence reports; AI-generated market simulations; competitor mapping; and pivot records.

Milestone 3 Prototype and digital venture development -

Students build prototypes, develop minimum viable product (MVP)s , create digital solutions, test functionality, and gather user feedback. They need to showcase failed prototypes, abandoned concepts, unexpected discoveries, technological dead ends, and lessons learned.

Evidence to provide: prototype evolution history; version-control records; user-testing evidence; failure logs; design iterations; and AI-assisted development records.

Milestone 4 Industry readiness and productization -

The module leader creates a simulated hostile environment of venture stress test with external stakeholders attempting to break it: investors reject it; regulators challenge it; competitors attack it; customers abandon it; and AI systems identify vulnerabilities. Students must respond in real time. They should refine operational models, assess risks, develop scaling strategies, engage industry parterres, and test implementation.

Evidence to provide: industry feedback; stress-test outcomes; risk mitigation plans; regulatory readiness audits; and scalability models.

Milestone 5 Venture launch & technopreneurship -

Module leader creates a simulated scenario that the founders disappear. Another team, industry partner, AI agent, or future cohort inherits the venture. Students act as the inheriting group attempting to operate, improve and continue the venture. The assessment is evaluated on whether it survives without its creators. Students therefore need to launch ventures, transfer knowledge, create sustainability systems, build operational governance, and demonstrate market engagement.

Evidence to provide: venture launch evidence; customer adoption data; sustainability plans; ecosystem partnerships; successor handover documentation; impact metrics.

Final integrative assessment –

At the end of the project, students do not need to submit a report or business plan or presentation, but instead they will submit a very brief reflection on their own transformation in how their assumptions changed, how they responded to failure, how they adapted, how they led, how they used AI, and how they evolved as innovators and technopreneurs.

Recipe 2 The project that never ends

- **Purpose**

Future-oriented PBL model in which projects are designed as living, evolving ecosystems that

continue across multiple cohorts rather than ending when a module/programme finishes. Students inherits, critique, improve, extend and hand over evolving projects to future cohorts. Therefore, the project in nature becomes longitudinal, intergenerational, adaptive, and continuously ongoing.

It aims to position students as contributors to long-term knowledge ecosystems rather than temporary project assignment producers. It simulates authentic professional and research environments where work is inherited, revised and continuously improved and evolved. It encourages students to engage critically with previous work rather than always starting from scratch. It is also designed to develop students' ability to work within evolving AI-human systems.

It aims to shift learning from "completion" towards "continuous improvement and legacy building". Students are evaluated not only on what they create, but also on how effectively they improve what already exists and how valuable their contribution becomes for future cohorts.

- **Class Size**

50-400 students. 400+ students will require strong digital infrastructure and AI-supported project management.

- **Suitable Disciplines**

Science; Life sciences; Engineering, computer science and AI; Film, creative arts and technology; Design; Business, finance, economics and management; Humanities and social sciences; Education.

- **Assessment Examples**

Science as an example

Students participate in a continuously evolving scientific discovery programme investing previously unresolved biological or environmental systems. Each cohort inherits research data, experimental results, failed methods, unresolved anomalies, hypotheses, and partially answered scientific questions from previous students.

Students need to review previous cohort findings, identify methodological weakness, replicate or challenge earlier conclusions, redesign experiments, retest scientific hypotheses, generate new data, document unresolved questions, and prepare research handover for future cohorts.

Evidence from students for assessment may include: research audit reports including critique

of previous findings; experimental redesign logs which may cover improved methodologies; replication outcomes including validation or challenge of prior results; AI critique records involving evaluation of AI-generated hypotheses; scientific uncertainty maps which might identify unresolved problems; research continuity handover including guidance for future cohorts; reflective scientific reasoning statement with key understanding of science as evolving inquiry.

Recipe 3 Client-focused project deliverables

- **Purpose**

To transform theoretical problem-solving into authentic professional practice by connecting student work with real organizational needs and practical relevance.

- **Class Size**

30-60 students allows for meaningful client interaction and manageable project coordination while maintaining individual accountability.

- **Suitable Disciplines**

Business, Engineering, Architecture, Computer Science, Environmental Studies, and any applied professional fields.

- **Assessment Examples**

Consultancy Reports for Local Organizations: Small groups analyze real problems for community partners, presenting findings and recommendations. Assessment includes client feedback (20%), peer evaluation (20%), and instructor evaluation of final report (60%).

Design Prototypes with User Testing: Engineering/design students create prototypes for identified user needs, then conduct and analyze user testing sessions. Assessment values iteration based on user feedback.

Policy Briefs with Stakeholder Analysis: Students research complex social problems and produce policy recommendations considering multiple stakeholder perspectives, with assessment focused on practical feasibility and ethical considerations.

Recipe 4 PBL via collaborative platforms

- **Purpose**

To make problem-based learning feasible in large classes through structured online collaboration, combining the benefits of authentic problem-solving with scalable peer learning and assessment mechanisms.

- **Class Size**

80-300 students. This range is large enough to incorporate diverse teams and perspectives but structured to prevent individual disengagement.

- **Suitable Disciplines**

Education, Environmental Science, Humanities, and any field benefiting from diverse perspectives on complex problems.

- **Assessment Examples**

Crowdsourced Case Solutions: Large class divided into expert teams (research, analysis, solution design). Each team contributes to a shared solution wiki. Individual grades based on specific contributions + peer evaluations.
Parallel Project Teams: Multiple small teams work on variations of the same core problem. Assessment includes team deliverables (50%), cross-team peer reviews (25%), and individual reflections comparing approaches (25%).
Graduated Responsibility Projects: Complex problems broken into phases. Students rotate roles (researcher, analyst, synthesizer, presenter) across phases. Assessment combines phase-specific rubrics with overall project evaluation.

Recipe 5 Multi-agent AI project studio

- **Purpose**

To simulate real-world interdisciplinary collaboration by using AI agents as stakeholders, domain experts, critics, or users within project-based learning environments.

- **Class Size**

20–200 students

- **Suitable Disciplines**

Design, Engineering, Science

- **Assessment Examples**

1. AI urban planner + architect + environmental consultant simulation

Students engage with multiple AI "agents," each representing a different professional perspective. The task requires students to negotiate conflicting recommendations and develop an integrated proposal by balancing sustainability, functionality, aesthetics, and feasibility.

2. Smart city design studio with AI stakeholder negotiations

Students design urban interventions while interacting with AI-generated stakeholder personas such as residents, policymakers, businesses, and environmental groups. Assessment focuses on students' ability to manage competing priorities and justify decisions.

Recipe 6 Challenge-based community innovation project

- **Purpose**

To engage students in solving authentic societal or industry challenges through interdisciplinary collaboration, AI-supported research, and iterative prototyping.

- **Class Size**

30–300 students

- **Suitable Disciplines**

Business, finance, economics and management; Engineering; Design; Life Sciences; Science; Education

- **Assessment Examples**

1. Sustainable campus redesign challenge

Students investigate sustainability issues on campus and propose interventions informed by AI-supported data analysis and scenario modelling. Final outputs may include policy recommendations, design proposals, or implementation plans evaluated by internal stakeholders.

2. AI-supported community wellbeing proposal

Students work with local or simulated community challenges related to wellbeing, inclusion, or accessibility. AI tools support stakeholder analysis and idea generation, while assessment focuses on feasibility, empathy, and responsiveness to real community needs.

3. Smart mobility or urban resilience project

Students design solutions addressing urban congestion, disaster preparedness, or environmental resilience using AI-assisted simulations and forecasting tools. Assessment emphasises systems thinking, adaptability, and interdisciplinary integration.

Recipe 7 Multi-staged assessment model

- **Purpose**

To disaggregate traditional, high-stakes summative assessments into a progressive, learner-centered model. By breaking down major assignments into sequential, competence-based stages, this approach shifts focus from a single final product to the visible development of student competencies over time. This structure preserves student cognitive agency, integrates structured feedback loops, and mitigates academic integrity risks by making the iterative process of learning and the responsible use of AI explicit and verifiable.

- **Class Size**

20–60 students. Can be scaled to larger classes by using multiple instructors, shifting from written to oral/video feedback, providing selective instructor feedback, or appointing student champions to deliver structured peer feedback that complements instructor feedback.

- **Suitable Disciplines**

Humanities and social sciences; Education; Business; STEM; Data Science; Creative Arts and media

- **Assessment Examples**

Step 1: Review Your Module and Identify the Summative Assessment

Review your module individually or talk to colleagues with curriculum design expertise, and identify the major summative assessment of your module that aligns with your module aims and learning outcomes. You shall use this major summative assessment to develop your students core competences.

Step 2: Decompose the Summative Assessment

Decompose the summative assessment output into several logical, sequential milestones (e.g., for a 13-week module, you may plan 3-4 stages/milestones).

Step 3: Define the Human-AI Loop and Roles

Establish the Human-AI loop. For each milestone, define the distinct roles of the learner (e.g., apprentice, collaborator, project owner), the AI tool (e.g., helper, coach, partner), and the instructor (e.g., consultant, feedback provider, facilitator). Provide subject- and task-related critical AI-use guidance, such as prompts or strategies, to support students' critical and responsible use of AI.

Step 4: Execution and Refinement

Draft target timelines and sub-outputs. Create mini-deadlines across the semester. At each milestone, students must submit a specific intermediate artifact to keep work on track. For example, for a 13-week module, you may design weekly or bi-weekly project tasks aligned with the pre-defined 3-4 stages, which contribute directly to the summative assessment.

Step 5: Determine Feedback/Feedforward Mechanism and Schedule

Determine your feedback and feedforward approach and schedule. For example, you may schedule brief weekly or bi-weekly oral consultations during or after class (it is recommended to do so during class to ensure interaction and immediate feedback), peer feedback sessions, or presentation feedforward reviews at the end of each stage. Students should not progress to the next stage without completing the previous checkpoint. You may also use features or tools in the Learning Mall to facilitate this process, such as Level Up.

Step 6: Implement Critical Assessment and AI Reflections

Require students to document regular learning reflections and maintain an AI-use log or declaration, detailing how they prompt, critique, and verify generative AI outputs at each stage.

Step 7: Consolidate into a Final Portfolio

At the end of the module cycle, students synthesise their cumulative, revised stage outputs alongside an overall critical reflection for final grading. The assessment rubric should include process-based and evidence-based dimensions to guide students toward more ethical and effective use of AI.

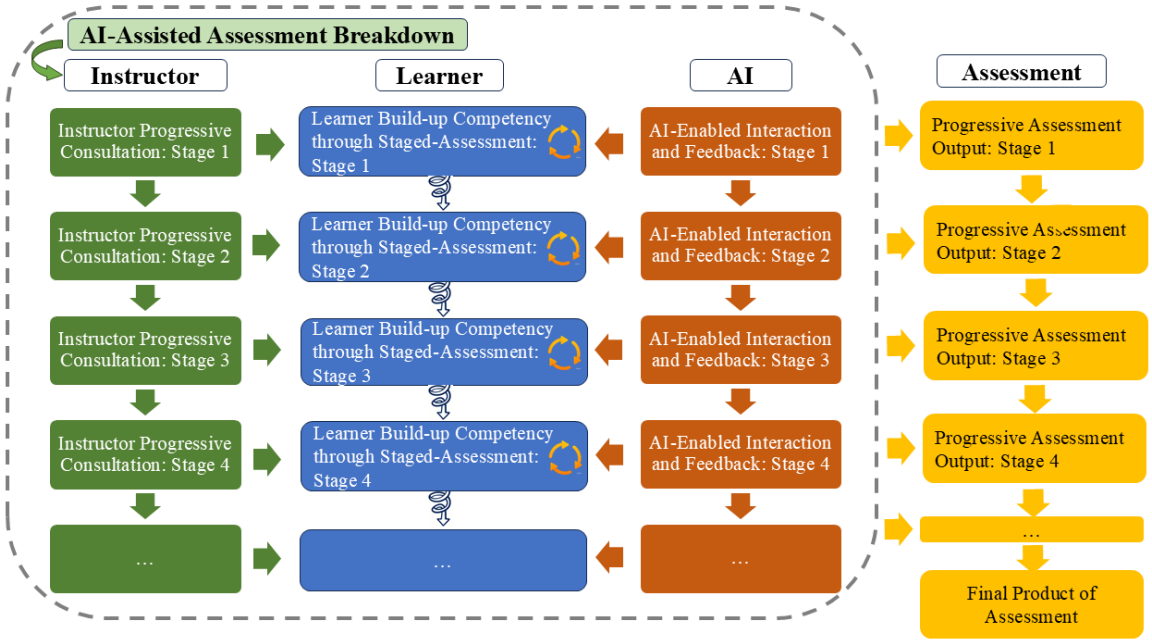


Figure 1. *The Multi-Staged AI-Enabled Assessment Model*

Recipe 8 AI-orchestrated learning intelligence framework

- **Purpose**

The AI-Orchestrated Learning Intelligence Framework is proposed as a future-oriented, conceptual extension of the Multi-Staged Assessment Model. While the Multi-Staged Model focuses on immediate pedagogical practice by manually restructuring assessments into progressive milestones, this framework outlines how integrated, AI-enabled capabilities can automate, connect, and scale these processes across the curriculum.

This framework is designed to transition institutional assessment from a static, submission-based, product-dominant model into an active, process-intelligent, developmental approach. Rather than evaluating only the final, static output, which generative AI can easily produce, this framework continuously captures, organizes, and visualizes a student's learning trajectory in real time. It prioritizes the evaluation of higher-order human competencies, critical AI literacy, and reflective decision-making across the entire learning journey.

- **Class Size**

50–300 students.

- **Suitable Disciplines**

All disciplines

- **Assessment Examples**

The Multi-Staged Model and the AI-Orchestrated Framework operate as a continuous pedagogical spectrum, representing different levels of institutional development:

The Multi-Staged Assessment Model (Current Practical Layer): Instructors manually design milestone structures, coordinate student-AI interactions, monitor progress, and provide manual feedforward within current LMS limitations and institutional regulations.

The AI-Orchestrated Learning Intelligence Framework (Future Evolution Layer): Intelligent, integrated AI systems automate workflow orchestration, dynamically synthesize learning evidence across modules, analyze collaborative contributions, and generate real-time feedback with traceable validation links.

Pillar One: Dynamic Process & Trajectory Profiling

This pillar redefines assessment as the evaluation of a student's developmental learning journey rather than isolated performance outcomes. By continuously analyzing learning trajectories, feedback engagement, and iterative progress, the framework captures authentic growth over time. Supported by an AI learning companion and portfolio-to-employment translation mechanisms, it enables personalized learning support while generating evidence-based competency profiles that are meaningful to both educators and future employers.

Pillar Two: Collaborative & Contribution Intelligence

This pillar enhances the fairness and transparency of group-work assessment by making individual contributions visible throughout the collaborative process. Through continuous contribution tracking and the triangulation of peer, self, instructor, and system-generated evidence, the framework recognizes both team outcomes and individual effort. This approach supports equitable grading, promotes accountability, and reduces the risk of free-riding in collaborative learning environments.

Pillar Three: Traceable Staff Intelligence Dashboards

This pillar empowers educators to manage process-oriented assessment efficiently while maintaining human oversight. Real-time dashboards provide actionable insights into student engagement, progress, and potential learning risks, enabling timely interventions. At the

same time, transparent AI-generated recommendations are linked to verifiable student evidence, ensuring that instructors can understand, validate, and retain ultimate responsibility for assessment decisions while reducing administrative workload.

Pillar Four: AI-Scaffolded Competency & Reflection

This pillar positions AI as a developmental partner that supports learning, reflection, and competency growth rather than merely producing outputs. By progressively scaffolding AI use from helper to coach and collaborator, the framework encourages students to develop critical AI literacy, reflective capacity, and responsible judgment. Assessment focuses on AI-resistant human competencies such as inquiry, critical evaluation, adaptability, and ethical decision-making, preparing learners for effective human-AI collaboration in future professional contexts.

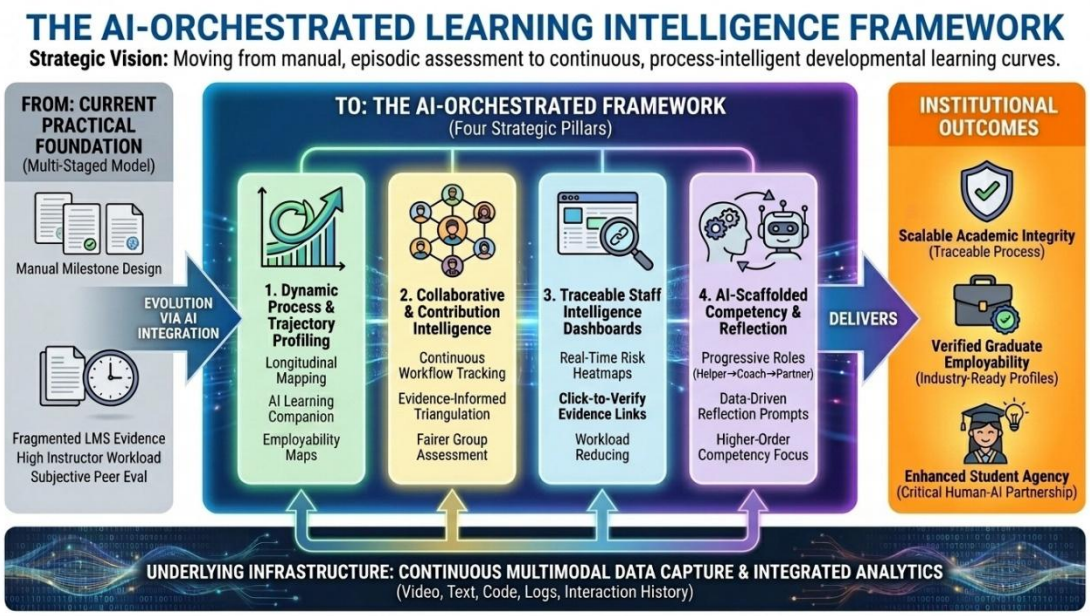


Figure 2 *The AI-Orchestrated Learning Intelligence Framework*

Section E: Assessment redesign implementation plan

Phase 1: Foundation & Capacity Building (September 2026-January 2027)

Goal: Create the enabling conditions, resources, and initial community of practice.

Key Actions

Establish a TAMA Steering Group: Form a cross-functional group with EDU leadership, 2-3 ADLTs, a student representative, and an AI technical expert. This group will champion and refine the framework.

Develop the "TAMA Toolkit" (Lead: EDU)

Rubric & Descriptor Bank: Create draft assessment rubrics for each competency (HI, AI, SI) at different performance levels aligned with XJTLU marking descriptor templates

Assessment Blueprints: Design 3-5 exemplar assessment tasks (e.g., a "Synergistic Portfolio" for a final-year project or an "AI Ethics Audit" for a business case).

Faculty Guide: A practical handbook on "Redesigning Your Assessment with TAMA."

Launch a "TAMA Pioneer Programme": Invite expressions of interest from faculty across schools to form a cohort of 15-20 early adopters. Offer them dedicated EDU support, small seed funding, and recognition.

Conduct Baseline & Needs Analysis: Survey faculty and students on current assessment practice and perceived gaps

Deliverables by January 2027

- TAMA Steering Group convened and terms of reference set.
- Version 1.0 of the TAMA Toolkit published internally.
- Pioneer Faculty Cohort identified and onboarded.
- Baseline survey data compiled and analyzed.

Phase 2: Piloting & Evidence Generation (January 2027-August 2027)

Goal: Test, refine, and generate proof-of-concept evidence from real modules.

Key Actions

Pilot Implementation: Pioneer faculty integrate TAMA-aligned assessments into their modules with a focus on combining 2 of the 3 dimensions (e.g., HI+AI in a reflective critique; AI+SI in a project workflow).

Intensive EDU Support: EDU assigns dedicated Educational Developers to partner with piloting faculty for co-design, rubric calibration, and marking moderation.

Formative Evaluation

- Faculty Feedback: Conduct structured interviews and focus groups with pioneers on workload, clarity, and student engagement.
- Student Feedback: Gather student perceptions via surveys and focus groups on the learning experience and perceived skill development.
- Workload Analysis: Monitor the time investment for marking and feedback compared to traditional assessments.

Showcase & Disseminate: Host a "TAMA Pilot Showcase" symposium in May 2027. Pioneers present their redesigned assessments, student work samples, and initial findings. Open invitation to all academic staff to build awareness and interest.

Deliverables by August 2027

- Piloted TAMA assessments across multiple schools.
- A comprehensive evaluation report with qualitative and initial quantitative data.
- Refined Version 2.0 of the TAMA Toolkit, incorporating lessons learned.
- Increased campus-wide awareness and curiosity about the framework.

Phase 3: Scaling & Integration (September 2027-January 2028)

Goal: Move from isolated pilots to school-level adoption and policy alignment.

Key Actions

School-Level Roadmaps: Work with ADLTs to identify 2-3 "signature programmes" per school where TAMA can be systematically integrated into the curriculum map across a programme

Expand Capacity Building: Launch open TAMA Design Workshops through each school, leveraging the pioneers as co-facilitators. Develop self-paced online modules.

Incentivise Adoption: Link TAMA redesign to existing mechanisms: make it a key component of Annual Review goals for teaching; include it as favourable evidence in promotion criteria; offer formal Micro-credentials for faculty who complete TAMA design projects.

Policy & System Alignment: Work with the Registry and EQA to update assessment policy language to recognise and encourage "human-AI collaborative" assessments and explore how TAMA rubrics can be integrated into the university's SCRP and UCRP review process for easier use.

Longitudinal Research Launch: Design a 2-3 year study to track the impact of TAMA

assessments on student outcomes (e.g., performance in later studies, graduate employment metrics, employer feedback).

Deliverables by January 2028

- School-specific implementation plans for 3+ major programmes.
- A university-wide faculty development programme on TAMA.
- Draft proposals for updated assessment policies and promotion criteria.
- A robust research plan for longitudinal impact assessment.
- TAMA firmly positioned as the institutional model for AI-era assessment at XJTLU.

References

Advance HE. (2025). *Framework for enhancing assessment*. Advance HE.

Australian Academic Integrity Network. (2023). *Generative artificial intelligence guidelines*. AAIN GenAI Working Group.

Australian Government Digital Transformation Agency. (2024). *Artificial intelligence impact assessment tool*. Digital Transformation Agency.

Australian Government Digital Transformation Agency. (2024). *Guidance for the artificial intelligence impact assessment tool*. Digital Transformation Agency.

Cardona, M. A., Rodríguez, R. J., & Ishmael, K. (2023). *Artificial intelligence and the future of teaching and learning: Insights and recommendations*. Office of Educational Technology, U.S. Department of Education.

Dollinger, M. & Nieminen, J. H. (2026). Reimagining Success and Failure: Equitable Assessment Practices in an Age of Artificial Intelligence. *Journal of University Teaching and Learning Practice*, 23(1). https://doi.org/10.53761/20z06b11

Elkhodr, M., et al. (2025). *Embedding generative AI in curriculum: The SAGE framework and evidence-based implementation guide*. Central Queensland University.

Fawns, T., Boud, D., & Dawson, P. (2026). Identifying what our students have learned: a framework for practical assessment validation. *Assessment & Evaluation in Higher Education*, 1–17. https://doi.org/10.1080/02602938.2026.2620053

Foltynek, T., Bjelobaba, S., Glendinning, I., Khan, Z. R., Santos, R., Pavletić, P., & Kravjar, J. (2023). ENAI recommendations on the ethical use of artificial intelligence in education. *International Journal for Educational Integrity, 19*(12). https://doi.org/10.1007/s40979-023-00133-4

Giugliano, A. (2025). *Generative AI policies at the world's top universities: October 2025*. The Sify. https://www.thesify.ai/blog/gen-ai-policies-update-2025

Higher Education Authority. (2025). *Generative AI in higher education teaching and learning policy framework*. Higher Education Authority, Ireland.

Kadel, R., Shailendra, S., & Saxena, U. R. (2025, September). Navigating the New Landscape: A Conceptual Model for Project-Based Assessment (PBA) in the Age of GenAI. In *2025 World Engineering Education Forum-Global Engineering Deans Council (WEEF-GEDC)* (pp. 1-9). IEEE.

Lodge, J., Howard, S., Bearman, M., Dawson, P., & Associates. (2023). *Assessment reform for the age of artificial intelligence*. Tertiary Education Quality and Standards Agency. https://www.teqsa.gov.au/sites/default/files/2023-09/assessment-reform-age-artificial-intelligence-discussion-paper.pdf

Miao, F. & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000386693

Quality Assurance Agency (2025). *How can generative AI be used in learning and teaching?* Quality Assurance Agency for Higher Education.

Russell Group. (2023). *Russell Group principles on the use of generative AI tools in education*. Russell Group. https://www.russellgroup.ac.uk/news/russell-group-hosts-dutch-universities-discuss-ai-education

Tertiary Education Quality and Standards Agency. (2025). *Enacting assessment reform in a time of artificial intelligence*. TEQSA. https://www.teqsa.gov.au/guides-resources/resources/corporate-publications/enacting-assessment-reform-time-artificial-intelligence

Tertiary Education Quality and Standards Agency. (2025). *Gen AI: Academic integrity and assessment reform*. TEQSA.

Weng, X., Xia, Q., Gu, M. Y. M., Rajaram, K., & Chiu, T. K. F. (2024). Assessment and learning outcomes for generative AI in higher education: A scoping review on current research status and trends. *Australasian Journal of Educational Technology, 40*(6), 37–55. https://doi.org/10.14742/ajet.9540

Xi, Y. (2025) Foreword: The Essence of Education in the AI Era. *Education + AI Strategic Framework 2025-2028.* https://www.xjtlu.edu.cn/wp-content/uploads/2025/03/XJTLU-Education-AI-Strategic-Framework.pdf

Xi'an Jiaotong-Liverpool University. (2025) *Education + AI Strategic Framework 2025-2028.* https://www.xjtlu.edu.cn/wp-content/uploads/2025/03/XJTLU-Education-AI-Strategic-Framework.pdf

Xi'an Jiaotong-Liverpool University. (2026) *Code of Practice on Assessments.*