**Framework for AI Task Automation in Higher Education**

**Introduction**

Higher Education Institutions (HEIs) are increasingly exploring Artificial Intelligence (AI) – including generative AI – to automate or assist with teaching, research, and administration. However, the rapid evolution of AI has outpaced policy in many institutions. A recent UNESCO survey found **fewer than 10% of universities have formal guidance on using generative AI**, leaving them vulnerable to risks like privacy breaches or inconsistent practices ([7 principles on responsible AI use in education | World Economic Forum](https://www.weforum.org/stories/2024/01/ai-guidance-school-responsible-use-in-education/#:~:text=,use%20of%20AI%20in%20education)). This updated taxonomy framework provides comprehensive guidance for evaluating where and how AI can be applied in HEI tasks. It integrates the latest global principles – notably the EU’s risk-based AI Act, UNESCO and OECD AI ethics guidelines – to ensure AI use is **responsible, ethical, and aligned with educational values**. The framework explicitly covers both academic tasks (teaching, learning, research) and professional services (administration, student support, etc.), including scenarios involving student-facing AI tools. Two outputs are provided: (1) a formal guidance document explaining the framework’s rationale, dimensions and categories, and (2) a practical checklist for staff to easily apply the framework in decision-making.

**Rationale and Context**

**Why an updated framework?** Universities face pressure to innovate with AI while safeguarding academic integrity, equity, and trust. Recent developments offer new guardrails for AI adoption. The **EU AI Act (2024)** introduces a risk-based approach, classifying AI systems from minimal to unacceptable risk ([What is the EU AI Act? A comprehensive overview](https://feedbackfruits.com/blog/from-regulation-to-innovation-what-the-eu-ai-act-means-for-edtech#:~:text=1,ensure%20their%20safety%20and%20reliability)). Education is explicitly deemed a **“high-risk” AI domain requiring strict compliance** ([What is the EU AI Act? A comprehensive overview](https://feedbackfruits.com/blog/from-regulation-to-innovation-what-the-eu-ai-act-means-for-edtech#:~:text=The%20implications%20of%20the%20EU,trajectories%20and%20their%20future%20careers)) ([What is the EU AI Act? A comprehensive overview](https://feedbackfruits.com/blog/from-regulation-to-innovation-what-the-eu-ai-act-means-for-edtech#:~:text=,detect%20prohibited%20behavior%20during%20assessments)) – for example, AI used in admissions or grading must meet rigorous safety, transparency and oversight standards. UNESCO’s *Recommendation on AI Ethics* (2021) and **guidance on AI in education** advocate a human-centric approach, emphasising that AI should **safeguard human rights and dignity, promote transparency, and avoid bias or discrimination (**[**A comprehensive AI policy education framework for university teaching and learning | International Journal of Educational Technology in Higher Education | Full Text**](https://educationaltechnologyjournal.springeropen.com/articles/10.1186/s41239-023-00408-3#:~:text=The%20UNESCO%20framework%20for%20AI,the%20following%20recommendations%20are%20provided)**)**. Similarly, the **OECD AI Principles** call for **trustworthy AI that respects human rights, fairness, and privacy**, with mechanisms for human oversight and accountability ([AI principles | OECD](https://www.oecd.org/en/topics/sub-issues/ai-principles.html#:~:text=Human%20rights%20and%20democratic%20values%2C,including%20fairness%20and%20privacy)) ([AI principles | OECD](https://www.oecd.org/en/topics/sub-issues/ai-principles.html#:~:text=To%20this%20end%2C%20AI%20actors,the%20state%20of%20the%20art)). These global guidelines, along with sector-specific insights (e.g. Jisc and EDUCAUSE resources), inform our framework to ensure it is robust, up-to-date and internationally aligned. By incorporating ethical principles and a risk-based lens, the framework helps HEIs balance **innovation with responsibility** – leveraging AI’s benefits (efficiency, personalisation, analytics) while mitigating potential harms (safety failures, bias, loss of human connection).

**Scope:** This framework covers the broad spectrum of university functions. It applies to academic **teaching and learning tasks** (e.g. content delivery, assessment, feedback), **research tasks** (literature review, data analysis, writing), and **professional services** such as admissions, registry, library, IT support, HR, finance, facilities management, and student services. Importantly, it includes **student-facing AI tools** – for instance, AI tutors, writing assistants, or chatbots that interact directly with students. The goal is to provide a consistent evaluative approach across all these contexts. Not every AI use is appropriate: some tasks demand a human touch or carry ethical risks that outweigh automation gains. This framework helps decision-makers determine the **appropriate level of AI involvement** – whether a task should remain human-led, can be AI-assisted, has high automation potential, or could be fully automated – based on a multi-dimensional assessment.

**Methodology for Framework Development**

To update the taxonomy, we combined literature review with expert input. We began with the six core dimensions of evaluation from the previously proposed framework and **refined each dimension against current best practices and policies**. Key references included the EU AI Act’s risk classification (to embed a risk assessment mindset), UNESCO and OECD AI ethics principles (to embed global values), and **higher education case studies** of AI implementation. Each proposed dimension was reviewed in light of these sources and expanded to cover both academic and administrative contexts. We ensured the framework aligns with the **EU’s risk tiers** (e.g. high-risk educational uses requiring human oversight ([What is the EU AI Act? A comprehensive overview](https://feedbackfruits.com/blog/from-regulation-to-innovation-what-the-eu-ai-act-means-for-edtech#:~:text=These%20applications%20must%20comply%20with,to%20maintain%20fairness%20and%20transparency))) and integrates ethical checkpoints (e.g. avoiding AI applications that violate UNESCO’s humanistic values ([A comprehensive AI policy education framework for university teaching and learning | International Journal of Educational Technology in Higher Education | Full Text](https://educationaltechnologyjournal.springeropen.com/articles/10.1186/s41239-023-00408-3#:~:text=The%20UNESCO%20framework%20for%20AI,the%20following%20recommendations%20are%20provided))). The updated framework was validated against diverse university scenarios – from automating routine clerical processes to using AI in pedagogical roles – to ensure its guidance is **nuanced yet practical**. Finally, two formats were developed: a detailed guidance document (below) for thorough understanding, and a condensed checklist for everyday use by staff evaluating AI solutions.

**Core Dimensions for Evaluating AI Use**

When considering AI automation or assistance for a given university task, staff should evaluate it along **six key dimensions**. These dimensions ensure a holistic understanding of the task, the AI’s fit, and the implications of its use. The dimensions are:

**1. Task Nature & Complexity**

**What is the nature of the task and how complex is it?** This dimension examines the characteristics of the task itself: its purpose, structure, and requirements. Key considerations include:

* **Type of Task:** Is it academic (e.g. teaching, grading, research) or administrative (e.g. scheduling, data entry, student support)? Understanding the context helps align with relevant policies (for instance, teaching-related tasks might invoke academic integrity considerations, whereas administrative tasks might focus on efficiency and data accuracy).
* **Complexity & Structure:** How complex or predictable is the task? Routine, well-defined tasks (like tallying attendance or processing payroll data) are more amenable to automation. In contrast, tasks requiring **creativity, critical thinking or situational judgment** (like designing curriculum or resolving a student grievance) have high complexity and are less straightforward for AI. The more nuanced and context-dependent a task, the more caution needed in applying AI.
* **Standardisation:** Does the task follow a clear procedure or set of rules, or does it involve case-by-case variation? Highly standardised tasks (e.g. formatting bibliographies) can often be handled by AI reliably. Tasks that are **open-ended or involve unstructured inputs** (e.g. an open-ended essay feedback or a research hypothesis generation) may be harder for AI to perform consistently well and likely need human interpretation.
* **Frequency & Scale:** Consider how often the task occurs and the volume of work. High-frequency repetitive tasks (such as sorting incoming student enquiries or updating records) might be good candidates for automation due to volume. Rare or one-off tasks might not justify the effort of developing an AI solution, unless they are very critical.
* **Knowledge & Domain Requirements:** Does the task require deep domain expertise, tacit knowledge, or contextual awareness that AI might lack? For example, academic advising requires understanding a student’s personal circumstances and aspirations – knowledge that goes beyond data and may be hard for AI to fully capture.

*Rationale:* Evaluating task nature and complexity helps identify *inherent limitations* for AI. Simpler, rule-based tasks align well with current AI capabilities, whereas highly complex or context-rich tasks highlight the continued importance of human expertise. This dimension prevents overreach – avoiding trying to automate what is not yet feasible or wise – and ensures that AI is applied where it can truly augment or improve processes.

**2. Risk, Safety & Security Assessment**

**What are the risks and safety implications of using AI for this task?** This dimension addresses potential harms and regulatory considerations, aligning with the **EU AI Act’s risk-based approach**. Key aspects include:

* **Impact and Consequence of Errors:** If an AI system makes a mistake in this task, what is the potential harm? Tasks that impact people’s futures or rights (e.g. admissions decisions or grading of exams) are **high-stakes**, where AI errors could seriously affect a student’s life. The EU’s draft AI Act labels such educational uses as *“high-risk”*, requiring strict oversight and accuracy ([What is the EU AI Act? A comprehensive overview](https://feedbackfruits.com/blog/from-regulation-to-innovation-what-the-eu-ai-act-means-for-edtech#:~:text=,detect%20prohibited%20behavior%20during%20assessments)). In these cases, any AI involvement must be very carefully managed (e.g. always with human review ([What is the EU AI Act? A comprehensive overview](https://feedbackfruits.com/blog/from-regulation-to-innovation-what-the-eu-ai-act-means-for-edtech#:~:text=These%20applications%20must%20comply%20with,to%20maintain%20fairness%20and%20transparency))) or perhaps avoided until proven safe. By contrast, low-stakes tasks (like auto-formatting a document) carry minimal risk if the AI errs.
* **Risk Level Classification:** Following the EU AI Act, consider if the AI application would be **Unacceptable, High, Limited, or Minimal risk** ([What is the EU AI Act? A comprehensive overview](https://feedbackfruits.com/blog/from-regulation-to-innovation-what-the-eu-ai-act-means-for-edtech#:~:text=1,ensure%20their%20safety%20and%20reliability)). *Unacceptable risk* (e.g. AI that breaches fundamental rights or is manipulative) should not be deployed at all in HEIs. For example, **emotion recognition AI on students is explicitly prohibited in EU educational contexts (**[**What is the EU AI Act? A comprehensive overview**](https://feedbackfruits.com/blog/from-regulation-to-innovation-what-the-eu-ai-act-means-for-edtech#:~:text=One%20of%20the%20most%20notable,adversely%20affect%20their%20educational%20experience)**)** due to privacy and manipulation concerns. *High-risk* uses (e.g. AI for student admissions, grading, or disciplinary monitoring) demand robust risk controls: rigorous testing, documentation, human oversight, and compliance checks ([What is the EU AI Act? A comprehensive overview](https://feedbackfruits.com/blog/from-regulation-to-innovation-what-the-eu-ai-act-means-for-edtech#:~:text=,detect%20prohibited%20behavior%20during%20assessments)). *Limited-risk* uses (e.g. an AI chatbot for FAQs) may be allowed but should include transparency (users informed they’re interacting with AI) ([What is the EU AI Act? A comprehensive overview](https://feedbackfruits.com/blog/from-regulation-to-innovation-what-the-eu-ai-act-means-for-edtech#:~:text=applications%20used%20in%20critical%20infrastructure%2C,These%20can%20be%20put%20on)). *Minimal-risk* uses (most administrative automations) have no special regulatory requirements but still warrant basic diligence.
* **Safety & Security:** Does using AI introduce safety or cybersecurity issues? For instance, an autonomous lab apparatus or campus security drone must be fail-safe to avoid physical harm. A student-facing chatbot must be secure against data leaks and adversarial inputs. Data security is paramount – tasks involving personal student/staff data must ensure **compliance with privacy laws (e.g. GDPR)**, secure handling of records, and protection against breaches. Any AI system should be evaluated for vulnerabilities that could be exploited, and mitigation measures (encryption, access controls, human monitoring) should be in place.
* **Reliability and Robustness:** How mature and reliable is the AI technology for this task? Newer generative AI models can sometimes produce inaccurate or unpredictable outputs (“AI hallucinations”), which is risky for critical tasks. The system should be **thoroughly tested** under expected conditions and stress cases. If an AI’s performance is inconsistent or error-prone in a given task scenario, the risk may outweigh the benefit. Robustness also means the AI can handle diversity in the input (e.g. different student dialects in an AI tutor) without failing or bias.
* **Fallback and Control:** Are there mechanisms for human intervention or shutdown if things go wrong? Even in largely automated processes, there should be safety nets (human-in-the-loop or human-on-the-loop). For example, if an AI flags a student for misconduct, a human should verify before any penalty. If an AI scheduling system starts creating conflicts due to a bug, staff should be able to quickly override. Ensuring *human control* is a key part of managing risk ([AI principles | OECD](https://www.oecd.org/en/topics/sub-issues/ai-principles.html#:~:text=To%20this%20end%2C%20AI%20actors,the%20state%20of%20the%20art)). No AI system should be a black box that cannot be audited or switched off in emergencies.

*Rationale:* This risk assessment dimension ensures HEIs do not deploy AI blindly. By mapping the task to risk levels and considering safety/security, the institution can comply with regulations and uphold its duty of care to students and staff. High-risk applications in education are to be treated with extreme caution – for instance, **any AI grading system must include human oversight to maintain fairness and transparency (**[**What is the EU AI Act? A comprehensive overview**](https://feedbackfruits.com/blog/from-regulation-to-innovation-what-the-eu-ai-act-means-for-edtech#:~:text=These%20applications%20must%20comply%20with,to%20maintain%20fairness%20and%20transparency)**)**. Lower-risk automation can proceed more freely but still needs basic safeguards. Ultimately, this dimension protects against unintended negative outcomes, ensuring AI use does not compromise safety, privacy or security.

**3. Ethical & Legal Considerations**

**Is it ethically appropriate to use AI for this task, and does it align with legal/ethical principles?** This dimension probes issues of fairness, accountability, transparency, and social impact, drawing on frameworks like UNESCO’s AI ethics and the OECD principles. Key points:

* **Fairness and Non-Discrimination:** Would AI treatment of the task introduce bias or unfair outcomes? AI systems can inadvertently perpetuate bias present in training data. In a university context, this is critical for tasks like admissions scoring, recruitment shortlisting, or even an AI tutor’s responsiveness. We must ensure the AI’s decisions do **not disadvantage any group** (e.g. not biased against certain demographics). UNESCO and OECD guidelines emphasise that AI must respect human rights and values like equality and non-discrimination ([AI principles | OECD](https://www.oecd.org/en/topics/sub-issues/ai-principles.html#:~:text=AI%20actors%20should%20respect%20the,protected%20by%20applicable%20international%20law)). If an AI tool cannot meet these standards (for instance, a scholarship selection algorithm that tends to favour one gender due to biased data), its use would be unethical. Mitigation could include bias audits of the AI, diverse training data, or simply keeping humans in charge of decisions.
* **Transparency and Explainability:** Ethical AI use requires that stakeholders understand when and how AI is involved. Staff, students or applicants should be informed if an automated system is part of a decision process (for example, letting students know an AI will be used to proctor an exam or that an AI chatbot is answering their query) ([What is the EU AI Act? A comprehensive overview](https://feedbackfruits.com/blog/from-regulation-to-innovation-what-the-eu-ai-act-means-for-edtech#:~:text=applications%20used%20in%20critical%20infrastructure%2C,These%20can%20be%20put%20on)). Moreover, decisions affecting individuals (admission outcomes, grading feedback) should be explainable – either the AI is inherently interpretable or a human can explain the rationale. An AI whose workings are too opaque for accountability might be unsuitable for consequential decisions. The OECD calls for **responsible disclosure about AI systems** so that people can challenge or understand outcomes ([AI principles | OECD](https://www.oecd.org/en/topics/sub-issues/ai-principles.html#:~:text=AI%20Actors%20should%20commit%20to,with%20the%20state%20of%20art)) ([AI principles | OECD](https://www.oecd.org/en/topics/sub-issues/ai-principles.html#:~:text=led%20to%20the%20prediction%2C%20content%2C,system%20to%20challenge%20its%20output)). In practice, this might mean choosing AI tools that provide reasoning (or using simpler rules-based AI) for high-impact tasks, and always giving individuals a path to appeal or question an AI-driven decision.
* **Accountability and Human Oversight:** Who is accountable if the AI makes a wrong decision? HEIs must ensure that introducing AI does not dilute accountability. Typically, the institution (and specific staff overseeing the process) remain responsible for outcomes, even if AI was involved. This ties closely to having human oversight: **AI should assist, not fully replace, human responsibility in education (**[**7 principles on responsible AI use in education | World Economic Forum**](https://www.weforum.org/stories/2024/01/ai-guidance-school-responsible-use-in-education/#:~:text=Any%20AI,responsibilities%20of%20educators%20and%20administrators)**)**. Many ethical frameworks insist on maintaining human agency – for example, the **World Economic Forum’s education AI principles state that AI should augment but *not replace* educators’ and administrators’ responsibilities (**[**7 principles on responsible AI use in education | World Economic Forum**](https://www.weforum.org/stories/2024/01/ai-guidance-school-responsible-use-in-education/#:~:text=Any%20AI,responsibilities%20of%20educators%20and%20administrators)**)**. Thus, for tasks with ethical or judgment components (disciplinary actions, grading, advising), the framework leans towards AI as a supporting tool with humans making final decisions. Clear governance should be in place: define who reviews AI outputs and how decisions are validated.
* **Privacy and Consent:** Does the AI use personal data, and if so, is it compliant with privacy laws and ethical data use? Tasks that involve student or staff personal data (academic records, health or counseling notes, etc.) are sensitive. Using AI might require sharing data with third-party services or aggregating data in new ways. Ensure explicit consent is obtained where appropriate, and that data minimisation principles are followed (only the necessary data is used). For instance, an AI advising system might use a student’s academic history to recommend courses – the student should know and consent to that use of their data. Privacy is also about how the AI is hosted – if it’s a cloud AI service, the university must vet the provider’s data protection measures. Adherence to frameworks like GDPR is mandatory in British and European contexts.
* **Alignment with Academic Values:** Consider if automating the task could undermine core academic or ethical values. For example, fully automating student feedback might save time, but if the feedback becomes generic and offers no human mentorship, it could conflict with the educational value of personalised guidance. Likewise, an AI that writes research papers would raise questions of academic integrity. UNESCO’s humanistic approach reminds us AI in education should **enhance human capabilities and uphold human dignity (**[**A comprehensive AI policy education framework for university teaching and learning | International Journal of Educational Technology in Higher Education | Full Text**](https://educationaltechnologyjournal.springeropen.com/articles/10.1186/s41239-023-00408-3#:~:text=The%20UNESCO%20framework%20for%20AI,the%20following%20recommendations%20are%20provided)**)**, not erode the teacher-student relationship or student’s own effort. When evaluating a task, ask: *Does AI help us do this in a way that is ethical and educationally sound?* If there is any risk of encouraging misconduct (e.g. students relying on AI to do their assignments) or reducing the richness of human learning, those ethical downsides must be weighed heavily. Policies (like honour codes or AI tool usage guidelines) may need to accompany the technical implementation.

*Rationale:* This dimension ensures AI use in HEIs aligns with **ethical norms, legal requirements, and the institution’s mission**. By explicitly checking for fairness, transparency, accountability, and respect for persons, we prevent scenarios where AI could cause injustice or damage trust. Given public concern over AI (e.g. bias in algorithms, “black box” decisions), proactively addressing these issues also helps maintain legitimacy and community acceptance of AI initiatives. Ultimately, if a proposed AI application cannot meet the university’s ethical standards (or society’s legal standards), it should be rethought or rejected, regardless of potential efficiency gains.

**4. Stakeholder Value & Impact**

**Who are the stakeholders of this task, and how will AI affect them?** This dimension broadens the analysis to consider the value delivered and the impact on all parties – students, staff, faculty, administrators, and even external partners. Key considerations:

* **Benefits and Value Added:** Identify what positive outcomes AI could bring to the task for each stakeholder. For students, could AI provide faster feedback, more personalised support, or greater access to services? For academic staff, does it reduce drudgery (e.g. automating repetitive grading tasks) and free time for higher-level work? For administrators, does it improve efficiency or accuracy of processes (e.g. fewer manual errors in data entry)? It’s important to ensure that AI integration is **purposeful and tied to clear educational or operational goals (**[**7 principles on responsible AI use in education | World Economic Forum**](https://www.weforum.org/stories/2024/01/ai-guidance-school-responsible-use-in-education/#:~:text=1,of%20AI%20to%20educational%20goals)**)**. For example, using an AI tutor should aim to improve student learning outcomes or engagement, not just to appear “innovative.” Weigh the potential improvements in quality, speed, scalability or cost-saving that AI might offer for the task.
* **Negative Impacts or Trade-offs:** Consider possible downsides for stakeholders. Could AI introduction lead to job redundancy or deskilling of staff? For instance, automating a scheduling task might raise staff fears about roles being diminished. Could it frustrate users? (E.g. a poorly designed student support chatbot might irritate students if it can’t answer complex questions.) Does it risk lowering the quality of the output? (E.g. AI-generated feedback might lack the nuanced insight a professor would provide.) It’s crucial to gauge whether the **stakeholders will accept and trust the AI solution**. Early involvement of staff and student representatives can surface concerns – perhaps staff worry about over-reliance on AI, or students worry about fairness. If the negative impacts outweigh the benefits (e.g. significant resistance, harm to user experience, or loss of a “human touch” that stakeholders value), then the case for AI in that task is weak.
* **Equity and Inclusion:** Will the AI solution serve all stakeholders effectively, or are there equity concerns? For example, if an AI tool requires the latest smartphones or high bandwidth, some students may be disadvantaged. If an AI advising system primarily supports English language and your campus has many ESL students, will they be equally served? Stakeholder impact must include an equity lens to ensure AI doesn’t inadvertently widen gaps. The **goal is to use AI in ways that *benefit all groups* and promote inclusive education (**[**7 principles on responsible AI use in education | World Economic Forum**](https://www.weforum.org/stories/2024/01/ai-guidance-school-responsible-use-in-education/#:~:text=AI%20should%20be%20employed%20purposefully,diverse%20learning%20needs%20and%20backgrounds)**)**, in line with the principle of equity in AI deployment.
* **Stakeholder Involvement and Communication:** A practical aspect is how stakeholders will be informed and involved. Introducing AI should come with training or orientation for staff who will work with it, and clear communication to students or end-users about what the AI does. Stakeholders should have a channel for feedback – e.g. a way to report problems or suggest improvements if the AI is not meeting their needs. Over time, evaluating the impact means collecting input: are students finding the AI-generated feedback useful? Are staff satisfied that the automated system is accurate? This ties into a continuous improvement ethos (which we revisit in the evaluation checklist).
* **Strategic Alignment:** Does the AI use align with the institution’s strategic goals and values? For instance, if a university prides itself on “personalised learning through close faculty-student interaction,” then replacing too much of that interaction with AI might conflict with its brand and promise. Conversely, if a strategic goal is innovation or digital transformation, certain AI projects might have added institutional value (being seen as a leader, attracting students interested in tech, etc.). Stakeholder impact isn’t just operational, but also reputational and cultural.

*Rationale:* By assessing stakeholder value and impact, we ensure the decision to use AI is *demand-driven* (solving a real problem or adding clear value) and *stakeholder-informed*. It’s not technology for technology’s sake. This dimension roots the framework in the practical reality of the university environment – any AI solution will ultimately be judged by how well it serves people. By foreseeing impacts, we can plan mitigations (e.g. re-skilling staff for higher-value roles if certain tasks are automated) and make sure the transition is positive. A task should only be automated if it genuinely improves or maintains the quality of service/education for stakeholders; otherwise, maintaining a human-led approach might be preferable.

**5. Human Element Requirement**

**How important is human involvement for this task, and what is the role of human expertise or empathy?** This dimension gauges to what extent the *“human touch”* is essential. Some tasks intrinsically require human cognition, creativity, or compassion that AI cannot replicate, whereas others are mechanical. Considerations include:

* **Need for Empathy or Personal Interaction:** Many educational and support tasks rely on emotional intelligence and empathy. For example, counseling a distressed student, mentoring a postgraduate, or delivering a nuanced lecture often require human connection and understanding. AI, no matter how advanced, **cannot authentically replicate empathy or build trust in the way a human can**. If the task’s success hinges on personal rapport, cultural understanding, or ethical judgment, the human element is critical. In such cases, AI, if used at all, should be in a very limited supportive capacity (e.g. providing background data to the human counselor, but never interacting with the student on its own in sensitive matters).
* **Requirement for Human Judgment and Expertise:** Does the task call for complex decision-making that benefits from human expertise, tacit knowledge, or ethical considerations? Academic judgments like grading an essay for originality and insight, or deciding whether a PhD thesis meets the novelty standard, involve subtleties that go beyond checklists – they require an expert’s discernment. Similarly, disciplinary decisions or accommodations for a student require a holistic view of context and fairness. Humans are better at handling exceptions and moral dilemmas. If a task is not rule-based and involves weighing competing values or interpreting ambiguous information, **human judgment should lead** (with AI perhaps providing data but not making final calls). Even the EU’s regulations imply human oversight on high-stakes decisions to uphold fundamental rights ([What is the EU AI Act? A comprehensive overview](https://feedbackfruits.com/blog/from-regulation-to-innovation-what-the-eu-ai-act-means-for-edtech#:~:text=These%20applications%20must%20comply%20with,to%20maintain%20fairness%20and%20transparency)).
* **Level of Autonomy Appropriate:** This relates to *human-in-the-loop*. Would full automation undermine the quality or integrity of the task? For instance, a fully automated degree audit might miss nuances (like a course substitution approved under special circumstances) that a human advisor would catch. Having a human in the loop can ensure these nuances are accounted for. On the other hand, tasks that have **little need for human creativity or compassion** (like compiling statistics for a report) might have a low human element requirement, meaning automation is more acceptable. One way to evaluate this is to ask: *if a student or stakeholder found out this task was done entirely by a machine, would they feel uncomfortable or less served?* If yes, it indicates an expectation of human element.
* **Public Trust and Expectations:** The human element is also about perception. In academia, there is an expectation of scholarly and personal engagement. If a university announced that AI, not professors, would answer all student questions or that algorithms decide all grades, there would likely be public concern. Maintaining trust may mean deliberately keeping humans at the forefront of certain processes. The **“human-led” category of tasks (see below) will include those where human presence is non-negotiable**. AI should be positioned as a tool that *augments* humans. As one school district analogy put it, AI is like a GPS guiding a driver – helpful for navigation but **“ultimate control remains with the human (educator or student)” (**[**7 principles on responsible AI use in education | World Economic Forum**](https://www.weforum.org/stories/2024/01/ai-guidance-school-responsible-use-in-education/#:~:text=By%20way%20of%20example%2C%20Peninsula,%E2%80%9D)**)**. This captures the ethos: technology as support, people as leaders.
* **Opportunities for Human Value-Addition:** Sometimes automating a task could strip away an aspect that was valuable for human development. For example, grading homework might be tedious for instructors, but reading student essays also gives instructors insight into how students are thinking, allowing them to adjust teaching. If an AI did all that grading, the instructor might lose touch with student progress. So we consider if human involvement in the task has secondary benefits (relationship-building, oversight of learning, etc.) that would be lost if automated. If so, perhaps the AI should only assist while humans remain in the loop to capture those benefits.

*Rationale:* This dimension ensures we don’t lose sight of the **human-centric nature of education**. Universities are not factories; their core value lies in human intellectual and social development. By identifying tasks that truly need human elements, we preserve the quality and ethos of educational and support processes. At the same time, recognising tasks with low human requirement helps us confidently automate those and redeploy human talent to areas where it matters more. The framework thereby promotes *human-AI collaboration*: AI does the repetitive or data-heavy lifting, humans provide oversight, empathy, and expert judgment where needed.

**6. Feasibility & Implementation**

**How feasible is it to implement AI for this task, and what resources or changes are required?** Even if a task is suitable in principle, practical constraints might affect the decision. This dimension addresses the logistical and technical implementation factors:

* **Availability of AI Solutions:** Determine if there are existing tools or systems that can perform the task, or if it would require custom development. For some tasks, mature off-the-shelf solutions exist (e.g. AI-driven scheduling software, plagiarism detection tools). For others, especially novel or complex tasks, the AI might only be available via emerging research or would need to be specially built. If only experimental or untested solutions exist, implementation risk is higher. Conversely, proven solutions with support might make adoption easier.
* **Technical Complexity and Integration:** Consider how the AI would integrate into current systems and workflows. Does it require major IT infrastructure changes or new data pipelines? For example, implementing an AI that predicts student dropouts might need data from academic records, attendance systems, and engagement metrics – pulling those together can be non-trivial. The institution’s IT maturity matters here (Jisc’s AI maturity model can help gauge this). If integration is complex, factor in the time and cost. Also check compatibility with existing platforms (e.g. can the AI tool plug into the learning management system or student information system?).
* **Cost and Resource Requirements:** AI implementation can involve costs – licensing software, purchasing hardware (if on-premise computation needed), or hiring expertise to maintain it. Is the expected benefit worth the cost? If a task automation saves a small amount of staff time but requires an expensive system, it may not be justified. Consider also the ongoing maintenance: AI systems may need regular retraining or updates (especially if data patterns change or if it’s a machine learning model that could drift). Budgeting for the full lifecycle is important. If an AI writing assistant is provided to students, for instance, can the university support it at scale (API costs) and maintain it as models evolve?
* **Staff Skills and Training:** Do staff have or can they gain the skills to use and oversee the AI? Successful AI adoption often requires training staff to interpret AI outputs, manage exceptions, or handle new workflows. For example, if librarians get an AI tool for recommending resources, they need to know how it works to trust its suggestions and to explain them to patrons. If implementing an AI requires data science expertise, does the unit have access to those skills? Sometimes, the feasibility is limited not by the AI tech but by human capacity to implement and sustain it. Including training sessions or hiring/consulting AI experts might be part of the plan.
* **Timeline and Change Management:** How quickly can this be implemented, and how will the change be introduced? A phased pilot approach may be wise for higher-risk or complex tasks – test the AI in a limited setting, evaluate, then scale up if successful ([A comprehensive AI policy education framework for university teaching and learning | International Journal of Educational Technology in Higher Education | Full Text](https://educationaltechnologyjournal.springeropen.com/articles/10.1186/s41239-023-00408-3#:~:text=Pilot%20testing%2C%20monitoring%20and%20evaluation%2C,or%20with%20specific%20learner%20populations)). This dimension also considers policy or process changes: does using AI in this task require updating regulations (e.g. exam rules if AI auto-grading is introduced), or communicating new guidelines to students? Implementation should be accompanied by appropriate change management: clear announcements, documentation, and feedback mechanisms during rollout.
* **Monitoring and Evaluation Plan:** Feasibility isn’t just about day-one deployment, but ongoing viability. Plan how the AI’s performance will be monitored and evaluated regularly. This overlaps with an ethos of continuous improvement (as highlighted by frameworks like the WEF’s principle of continuous evaluation ([7 principles on responsible AI use in education | World Economic Forum](https://www.weforum.org/stories/2024/01/ai-guidance-school-responsible-use-in-education/#:~:text=7,impact%20of%20AI))). Having metrics or KPIs for the AI-enabled process (e.g. accuracy rate, turnaround time, user satisfaction) and a schedule for review can ensure the implementation remains on track and issues are caught early.

*Rationale:* This final dimension grounds the framework in reality. A theoretically great AI idea must also be deliverable. By examining feasibility and needed resources, we avoid scenarios where plans fail due to underestimating complexity or cost. It encourages a pilot-and-evaluate mentality, which is especially important given the fast-changing AI landscape – what’s cutting-edge today might be commonplace (or obsolete) in a few years. The framework thus remains practical: recommending AI use only where the institution can responsibly deploy and maintain it. In some cases, the feasibility check might delay a project (e.g. wait a year until a trusted vendor releases a suitable tool, rather than building a risky custom system now). In others, it will highlight requirements to address (budget approval, staff training) for successful implementation.

**Decision Categories for AI Integration**

Using the above dimensions, HEIs can map each task to one of four recommended **decision categories** regarding AI involvement. These categories range from keeping the task entirely human-led to fully automating it. The placement depends on the overall evaluation – considering task complexity, risk, ethics, stakeholder impact, human element, and feasibility together. In practice, many tasks will fall in a grey area between categories; the aim is to identify the *best-fit category* given current conditions, acknowledging that this can be revisited as technology and context evolve. The categories are:

**Category 1: Human-Led (Minimal AI Involvement)**

**Definition:** Tasks in this category should remain **predominantly human-driven**, with at most minimal assistance from AI (if any). These are tasks where human expertise, judgment or empathy is essential, and where automation could pose high risks or ethical issues. Often these involve high complexity, high stakes, or core educational values such that AI is only used in a subordinate role (e.g. as a reference tool), or not at all.

**Characteristics:** Human-led tasks typically score high on complexity, ethical sensitivity, or need for human element in our dimensions. They may also be high-risk tasks under the EU AI Act that mandate human oversight. The mantra here is “**AI can advise, but humans decide**.”

**Examples of Human-Led Tasks:**

* **Student Pastoral Care and Counseling:** Providing mental health support or personal counseling to students should remain human-led. While AI chatbots exist for mental health, universities would use them cautiously; the **empathy and trust of a human counselor is irreplaceable**. AI might assist by, for example, alerting counselors to students’ stress signals (if ethically permissible) or providing suggested resources, but conversations and care plans should be human-to-human.
* **Admissions Decisions (Final Review):** Deciding which applicants to admit into a programme, especially in competitive contexts, should be led by admissions tutors/committees. Given this is a high-stakes, high-risk decision affecting individuals’ futures ([What is the EU AI Act? A comprehensive overview](https://feedbackfruits.com/blog/from-regulation-to-innovation-what-the-eu-ai-act-means-for-edtech#:~:text=,detect%20prohibited%20behavior%20during%20assessments)), human judgment is crucial to interpret qualitative factors and ensure fairness. AI can help by processing large volumes of applications or flagging certain patterns (e.g. identifying if any required field is missing), but the ultimate decision, especially on borderline cases or where context matters, stays with humans. Humans can consider personal statements, references, and life circumstances in a holistic way that an algorithm might not.
* **Academic Integrity Investigations:** Determining if a student committed plagiarism or cheating, and what the penalty should be, involves ethical judgment and often a need for discussion with the student. AI can detect text similarities or unusual patterns (e.g. via plagiarism detection software or exam monitoring tools), but those are just inputs to a human-led investigation. The outcome (e.g. upholding an academic offence, issuing warnings) must be decided by faculty panels or honor boards, to ensure due process and consideration of context (perhaps the student had a misunderstanding or there were special circumstances). This maintains fairness and the educational aspect of such processes.
* **Curriculum Design and Pedagogy:** Designing course content, choosing teaching methods, and defining learning outcomes are creative academic tasks that require pedagogical expertise. AI tools can suggest content (e.g. generate quiz questions, provide data on skill gaps), but faculty should lead the design to align with educational philosophy and student needs. The **principle of “Pedagogy Before Technology” is key – human educators decide the goals and structure, and they use AI only insofar as it serves those goals (**[**What is the EU AI Act? A comprehensive overview**](https://feedbackfruits.com/blog/from-regulation-to-innovation-what-the-eu-ai-act-means-for-edtech#:~:text=Our%20design%20principles%20focus%20on,the%20following%20key%20areas)**)**. This ensures AI doesn’t dictate what or how we teach, but rather supports the educator’s vision.
* **High-Level Policy and Strategy Decisions:** University leadership decisions (such as creating a new degree programme, allocating budgets, or forming partnerships) remain human-led. AI analytics might inform these decisions (trends, forecasts), but given the complexity, multiple objectives, and moral accountability of such decisions, they cannot be handed off to automation. Leaders must weigh values, institutional mission, and stakeholder politics in ways AI cannot.
* **Individual Student Assessments Requiring Complex Judgment:** For example, grading a master’s thesis, conducting a viva voce (oral exam), or giving feedback on creative projects. These tasks involve nuanced understanding, academic judgment, and often interactive dialogue – things best done by academics. AI may help check factual accuracy or grammar in a thesis, but evaluating the contribution to knowledge and the quality of argument is a human scholarly responsibility.

These examples span academic, support, and administrative spheres, but share a theme: **the human is at the centre**. By keeping these tasks human-led, universities uphold ethical standards and quality. AI plays a secondary role if at all – e.g. a tool for information gathering or a second opinion – but never the primary actor. This category often correlates with tasks that are high risk or require trust; it aligns with the idea that some decisions (especially those affecting people deeply) must remain accountable to human professionals.

**Category 2: AI-Assisted (Human Primary, AI Supporting)**

**Definition:** In AI-assisted tasks, humans still lead and hold responsibility, but **AI is actively used as a tool to augment the human’s work**. The task is completed through a human-AI partnership: the AI handles certain sub-tasks or provides suggestions, and a human reviews, guides, or finalises the output. This category is about **enhancing human efficiency or capability** without removing the human from the loop.

**Characteristics:** Tasks suitable for AI assistance are those where AI can handle parts of the workload (especially repetitive or analytical parts) but human oversight is needed for quality, context, or ethical reasons. These tasks often have moderate complexity or stakes – enough that fully handing to AI is unwise, but with elements that AI can competently support. The risk is managed by continuous human involvement.

**Examples of AI-Assisted Tasks:**

* **Essay Grading with AI Feedback:** An instructor uses an AI tool to help grade essays or assignments. For instance, the AI might scan essays for grammar, give a preliminary score or highlight sections that seem off-topic or potentially plagiarised. The teacher then reviews those AI insights, reads the essay (perhaps more efficiently by focusing on flagged areas), and assigns the final grade with personalised comments. Here, AI speeds up the drudgery (spelling checks, finding where criteria may or may not be met) but **the educator remains the ultimate grader**, ensuring fairness and nuance in evaluation. This approach follows emerging practice – AI as a “second reader” that a human can agree or disagree with, not an automatic grader.
* **Student Support Chatbot (with Human Backup):** The university deploys an AI chatbot to answer common student queries (e.g. “When is the library open?”, “How do I reset my password?”). The AI handles queries instantly 24/7. However, it’s configured such that if it cannot confidently answer or if the question is complex (“I’m struggling academically, what should I do?”), it seamlessly hands off to a human staff member. Staff also monitor the chatbot logs and update its knowledge base as needed. In this AI-assisted scenario, **AI improves response time and offloads simple FAQs**, while humans handle exceptions and ensure the information stays correct. The student benefits from quick answers but isn’t left frustrated because human help is there for unusual or sensitive issues.
* **Admissions Pre-screening:** While final admission decisions are human-led (as noted above), AI can assist in pre-screening large applicant pools. For example, an AI system might automatically verify entry criteria (grades, prerequisites) and flag applications that meet certain benchmarks for further review, or highlight anomalies (such as a drastic discrepancy between test scores and grades) for closer attention. Admissions officers then focus their time on the substantive comparisons and interviews. The AI essentially does an initial triage. This speeds up processing and helps ensure no detail is overlooked, but crucially **every AI recommendation is reviewed by a person**, and officers can overturn or interpret the AI flags with full context.
* **Plagiarism Detection and Academic Integrity:** As mentioned, determining outcomes is human-led, but AI assistance is standard for detection. Tools like Turnitin (which use AI to detect similarity) are employed by instructors to catch potential plagiarism. The AI produces a report highlighting parts of a student’s submission that match other sources. The instructor then examines those highlights to judge if it’s plagiarism, a citation issue, or a false alarm. The AI assists by doing the heavy scanning against millions of sources – something impractical for a human – but the disciplinary decision remains with faculty. This combination respects due process and helps maintain high integrity standards efficiently.
* **Research Data Analysis (Human-Guided):** A research team might use AI/machine learning to analyze large datasets (for example, an AI finds patterns in a dataset of student performance, or in lab experiment results). The AI can detect correlations or clusters much faster than manual analysis. However, researchers remain in the loop to verify if those patterns are meaningful, to design the analysis parameters, and to interpret the results. They might iterate: run the AI analysis, see results, adjust the model or data cleaning, run again, etc. This is AI-assisted research – accelerating discovery but reliant on human scientific judgment to avoid false findings.
* **Content Creation with Human Editing:** This could apply to creating study materials or marketing content. For example, a staff member uses a generative AI to draft an outline for a lecture or to create a first draft of a university blog post. This draft saves time, but the staff member then **edits and fact-checks the AI-generated content** thoroughly, adding the needed human touch (ensuring tone and factual accuracy). The final output is thus higher quality and produced faster than the human alone might manage, yet it remains vetted and tailored by the human. This assists staff creativity and productivity without fully delegating it to AI.

In all these cases, the task’s quality and compliance are maintained by human oversight, but AI provides significant help – whether by shortening response times, processing large volumes of data, or offering preliminary results. **AI-assisted category is about leveraging AI’s strengths (speed, scale, pattern recognition) while keeping humans in charge of decisions and interpretations**. It’s a balanced approach that often yields efficiency gains and can improve consistency (AI is good at applying the same criteria uniformly), while humans handle the subjective or exceptional aspects. Many current successful applications of AI in universities fall in this category, as it navigates between innovation and caution.

**Category 3: High Automation Potential (Human Oversight on Mostly-Automated Task)**

**Definition:** Tasks in this category are **strong candidates for extensive automation**, where AI could perform the bulk of the work. Human involvement becomes more periodic or supervisory rather than continuous. The task can largely run on autopilot under AI control, *but with humans setting it up, monitoring outcomes, and handling edge cases or governance*. Essentially, the process is automated end-to-end for normal cases, and humans step in only for exceptions or to review aggregate performance.

**Characteristics:** Such tasks are typically low to moderate risk and relatively structured, so that AI can handle them reliably. They may not require real-time human input, and decisions made by the AI are either low-stakes or can be reversed/adjusted easily if needed. Importantly, even if the AI is doing most of the work, there is a **human-on-the-loop**: someone responsible for overseeing the system’s functioning, maintaining it, and intervening if anomalies occur. This category often emerges after positive experience in AI-assisted mode – once trust in the AI’s accuracy is established, one can move more responsibility to the AI.

**Examples of High Automation Potential Tasks:**

* **Timetabling and Course Scheduling:** Scheduling courses, rooms, and exam timetables is a complex logistical task that AI algorithms (optimisation solvers) are quite adept at. A modern university might use an AI-driven scheduling system to automatically assign class times and rooms based on input constraints (student course selections, room capacities, professor preferences). This system could generate an entire semester’s timetable with minimal human input after initial rules are set. **Staff would then review the draft schedule** for any glaring issues or special cases the AI couldn’t know (e.g. a one-time event needing a room), make minor adjustments, and then approve it. Over time, as the AI improves with feedback, the scheduling might become almost entirely automated except for final sign-off. This task, while complex, is rule-bound and doesn’t deeply affect individuals’ rights (aside from convenience), making it a safe candidate for high automation.
* **Library Chatbot for FAQs:** Extending the earlier chatbot example – if the AI assistant for library or IT support queries proves highly accurate, the university might allow it to function with minimal human monitoring. The bot could handle say 90% of questions on its own. Librarians or IT staff would only handle those few queries the bot escalates or any user complaints. Periodically, staff might review logs to retrain the bot or update information, but they are not actively answering each query. The majority of the workflow (question asked → answer provided) is automated. Given these are routine queries, the risk is low; the benefit is freeing staff from answering the same questions repeatedly. Humans remain available as a safety net, but the **AI operates with a high degree of autonomy day-to-day**.
* **Automated Attendance Monitoring:** Some institutions use AI-based systems (sometimes with computer vision or scanning) to record class attendance or track student presence. For example, students tap an ID card or an AI camera recognizes faces to mark attendance. The system then automatically compiles attendance records and can even send alerts if a student’s attendance drops (possibly indicating disengagement). In a high automation setup, this runs continuously with no need for manual roll calls. Staff involvement is only to check on alerted cases or maintain the system. Of course, such a system raises privacy issues (especially facial recognition), so an institution must weigh ethics and perhaps opt for less intrusive methods (cards or app check-ins). But from a process view, attendance tracking can be mostly automated. Staff advisors or tutors then **use the automated reports** to intervene with students who have issues – a good human-AI collaboration where AI does monitoring, humans do the mentoring.
* **Financial Transactions and Payroll Processing:** Administrative tasks like processing invoices, expense claims, or monthly payroll can be largely automated with AI and robotic process automation (RPA). For instance, an AI system could read invoice PDFs (using OCR), extract amounts and vendor info, match it to purchase orders, and flag any discrepancies or automatically schedule payment if everything matches. Finance staff then only handle exceptions (like an invoice that doesn’t match any record, or a flagged possible overcharge). Similarly, payroll software can automatically calculate salaries, taxes, etc., and deposit pay, with HR staff just reviewing summary reports. These tasks have clear rules and well-defined data, making them ripe for automation. The risk of error is not zero (financial errors matter), so oversight is needed – e.g. an accountant reviews a random sample or all high-value payments. But generally, the system runs with minimal intervention, improving efficiency and reducing manual errors.
* **Routine Communications and Notifications:** Universities often send out routine communications (e.g. deadline reminders, event announcements, feedback acknowledgments). An AI system can manage these at scale – for example, automatically emailing students who haven’t submitted coursework by a certain date with a polite reminder, or sending personalised nudges to students who are falling behind based on LMS data. Because the content of these messages can be templatized and triggered by defined conditions, AI (or automated scripts) can handle it without a person writing each message. Staff oversight might be just to ensure the templates are correct and to handle any replies that require a human. Over time, if such notifications prove beneficial (say increasing on-time submissions), they become a standard automated service. The impact on stakeholders is generally positive as long as messages are well-crafted (possibly even AI can tailor tone per student, though that enters nuanced territory).

In this category, **automation is high but not absolute**. The process can run on its own during normal operations, yielding significant efficiency gains, but humans haven’t completely “let go of the wheel.” They are supervising and ready to take control if needed (much like an autopilot with a pilot on standby). This ensures that if something goes outside expected parameters, human judgment can correct course. Many back-office processes in universities can reach this level – once initial trust in the AI’s accuracy is established, human role shifts to monitoring and improving the system rather than performing each transaction.

**Category 4: Full Automation (Autonomous AI with Oversight Checks)**

**Definition:** Full automation means the task is **handled entirely by AI or an automated system from start to finish**, with no routine human involvement in individual task instances. Human involvement is limited to **initial setup and periodic oversight or auditing**, rather than day-to-day execution. This category is reserved for tasks that are low-risk, highly structured, and where automation has proven to meet or exceed human performance reliably.

**Characteristics:** Fully automated tasks are typically those that are simple, repetitive, and where the cost of an error is very low (or easily mitigated by subsequent processes). These tasks likely scored very favorably on the feasibility dimension and posed minimal ethical or risk concerns. Even though the AI operates autonomously, the institution should still perform *periodic audits or maintenance* – “full automation” is not “forget about it entirely,” but it means no human choice is involved in each cycle of the task. It’s important that even fully automated processes have an owner who reviews them periodically for correctness and fairness (consistent with a principle of continuous evaluation ([7 principles on responsible AI use in education | World Economic Forum](https://www.weforum.org/stories/2024/01/ai-guidance-school-responsible-use-in-education/#:~:text=7,impact%20of%20AI))).

**Examples of Fully Automated Tasks:**

* **Auto-Grading of Objective Quizzes:** Grading of multiple-choice quizzes, fill-in-the-blank tests, or other objectively scored assessments can be fully automated. Many learning management systems already automatically grade quiz submissions and even provide immediate feedback. Since the answers are either correct or incorrect based on a key, AI can handle it without human graders. The only human role is in creating the quiz and answer key, and perhaps reviewing item statistics afterwards to see if any question was problematic. This is low risk – if a question is graded incorrectly due to a key error, instructors can always adjust scores after the fact (and such adjustments are easy to make upon review). Immediate, automated scoring benefits students (instant results) and saves instructor time, with essentially zero downside when properly set up.
* **Plagiarism Scan on Submission – Initial Stage:** While the interpretation of plagiarism reports is human-led (as discussed), the *act of scanning all submissions against databases* is fully automated. The system checks every paper uploaded to the VLE (Virtual Learning Environment) against millions of sources without any person initiating each check. The output (similarity score, highlighted matches) is then ready for a human to consider. The scanning task itself, however, runs autonomously for each submission. This automation ensures consistency (every single paper is checked) and operates in the background. It’s now a default part of many universities’ submission process – a classic example of an automated task embedded in workflow.
* **Routine Data Backup and Reports:** IT tasks like backing up servers or generating usage reports are fully automated via scripts or AI ops tools. For instance, every night an automated process might compile a report on LMS activity (logins, pages accessed) and email it to relevant staff, or simply archive it for records. Unless an anomaly occurs (no report received, indicating a failure), humans don’t intervene. Similarly, academic departments might have an AI script that compiles a draft of exam statistics or student progression data after each exam cycle, ready for the exam board to review. The assembly of the data is done without human labor. The academic staff then use that report to make decisions, but the task of generating the report required no manual effort each time.
* **Building Energy Management:** Many universities use smart building management systems driven by AI to control heating, cooling, and lighting to optimise energy usage. These systems adjust thermostats and lights based on occupancy data or schedules, entirely automatically. Facility managers set the parameters, but day-to-day adjustments happen via AI algorithms reacting to sensor inputs. This reduces energy waste and costs. As long as comfort thresholds are appropriately set, the impact on occupants is positive or neutral. Facility staff typically just audit energy reports and handle malfunctions, but don’t need to flip switches daily – a fully automated operational task.
* **Spam Filtering and Email Routing:** The university’s email system likely uses AI/ML filters to automatically detect spam or malicious emails and block/quarantine them without any human reviewing each message. Only if a user appeals (marking “this isn’t spam”) might an IT staff intervene in a specific case. This is a background process that protects security autonomously. It’s low-risk to automate because the cost of a mistake (a legitimate email mis-flagged) is minor and correctable, whereas reviewing millions of emails by hand would be impossible. Over time the AI models are retrained on feedback, but again this is often automated model updates by the email service provider. This illustrates how certain IT security tasks are best left to AI entirely.

In fully automated tasks, **the efficiency gains are maximised** – the process can run 24/7, at scale, with near-instantaneous handling and without requiring staff attention. This liberates staff from drudge work and allows them to focus on tasks from the other categories where their attention is more valuable. However, it’s crucial that HEIs still **audit outcomes and maintain accountability**. For instance, even if exam marking for MCQs is automated, the department should occasionally verify that the system graded correctly (no errors in answer key or system logic) – a simple check but important for due diligence. Similarly, for AI-driven processes like building management, monitoring KPIs (like energy saved vs. any complaints from occupants) ensures the automation is doing its job well. Full automation is not a “fire and forget”; it’s an achieved state of minimal intervention, kept on track by oversight mechanisms.

**Using the Framework and Acknowledging Nuance**

The above categories with examples show a range of AI integration in university tasks. It’s important to note that these are **not rigid boxes** – a task might move from one category to another over time, or straddle two categories depending on implementation. For example, an AI tutoring system might start as AI-assisted (with a teacher monitoring all interactions) and later become high automation once trust is built and the system improves, though it might never become fully autonomous due to the need for human pedagogical oversight. Context matters: two universities might categorize the same task differently based on their risk appetite, values, or resources.

The framework’s strength is in prompting careful evaluation. By going through the six dimensions, staff can justify *why* a task is kept human-led or is suitable for automation. It also encourages a **gradual approach** – perhaps trial an AI in an assisted capacity, evaluate stakeholder feedback (did it maintain quality? was it ethical?), then decide if further automation is warranted or not. This incremental adoption aligns with best practices (pilot testing and evaluation before full deployment ([A comprehensive AI policy education framework for university teaching and learning | International Journal of Educational Technology in Higher Education | Full Text](https://educationaltechnologyjournal.springeropen.com/articles/10.1186/s41239-023-00408-3#:~:text=Pilot%20testing%2C%20monitoring%20and%20evaluation%2C,or%20with%20specific%20learner%20populations))).

In summary, the updated taxonomy framework is designed to be **robust yet practical**. It acknowledges the complexity of university functions and the myriad factors to weigh (from EU regulatory compliance to the humane aspects of education), without being so overwhelming that nothing gets done. By structuring decision-making into clear dimensions and categories, it helps demystify the process of integrating AI. Universities can use it to ensure they embrace AI opportunities – improving services and efficiency – while **upholding ethics, safeguarding stakeholders, and preserving the fundamental human core of education**.

**Practical Checklist for Evaluating AI Use in University Tasks**

*Use this checklist to quickly assess a given task and decide on the appropriate level of AI involvement. For each dimension, answer the guiding questions to identify concerns or requirements. Then refer to the decision categories to determine if the task should be human-led, AI-assisted, highly automated, or fully automated. This checklist is meant to be a handy tool for staff and decision-makers to apply the framework in practice.*

**1. Task Nature & Complexity**

* **Type of Task:** Is the task academic (teaching, learning, research) or administrative (service, operations)? \_\_\_\_\_\_\_\_
* **Complex vs Routine:** Is it a routine, well-defined process or does it require creative, on-the-spot thinking? List any aspects that are highly complex or variable: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* **Structured Inputs/Outputs:** Does the task follow clear rules and criteria (structured), or does it involve unstructured data or open-ended outcomes? \_\_\_\_\_\_\_\_
* **Frequency/Scale:** How often does it occur and how many instances? (e.g. daily dozens of requests vs. annual review) \_\_\_\_\_\_\_\_
* **Expertise Needed:** Does it demand deep subject expertise or contextual knowledge? If yes, note what expertise: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*If the task is highly complex, unique, or requires a lot of creativity/human knowledge, lean towards keeping it human-led or only lightly assisted. Routine and high-volume tasks are better candidates for automation.*

**2. Risk, Safety & Security**

* **Consequence of Error:** What’s the worst-case scenario if AI makes a mistake in this task? (e.g. minor inconvenience, data loss, unjust outcome for a student) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* **Regulatory Risk Level:** Would this AI use be high-risk under regulations? (Does it affect student rights, safety, or critical decisions? E.g. admissions, grading = High risk; chatbot FAQ = Limited risk) \_\_\_\_\_\_\_\_
* **Data Privacy:** Does the task involve personal or sensitive data? If yes, ensure GDPR compliance and data protection measures. Any privacy flags? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* **Security:** Could automation introduce cybersecurity issues? (e.g. opening network access, vulnerability to hacking) List any security needs: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* **Safety:** For physical or well-being related tasks, can the AI operate safely? Any potential harm to people or infrastructure? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* **Human Control:** Will a human be able to intervene or override if the AI behaves unexpectedly? Plan for fallback: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*If the task is high-stakes or high-risk, ensure human oversight is built in and that it meets all compliance requirements. High risk might keep it in Human-Led or AI-Assisted. Low-risk tasks with good safety can be candidates for High Automation or Full Automation.*

**3. Ethical Considerations**

* **Fairness:** Could AI decisions in this task be biased or unfair to any group? How will you check and mitigate bias? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* **Transparency:** Do people (staff/students) need to know when AI is used here? How will you inform them? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* **Explainability:** Can the AI’s output/decision be explained to those affected? If not, should a human review it before finalising? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* **Accountability:** Who is accountable for outcomes of this task with AI? Has it been clearly assigned (e.g. a responsible staff or committee)? \_\_\_\_\_\_\_\_
* **Ethical Red Lines:** Does the AI use align with institutional values and codes of ethics? (For example, no surveillance beyond policy, no violation of academic integrity norms.) Any ethical concerns noted: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* **Consent:** If using personal data or affecting individuals (students/staff), do we need their consent or to offer an opt-out? Have we addressed this? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*Ethical red flags (bias, opacity, rights violations) usually mean either redesigning the AI approach or keeping a strong human role. If ethics check is satisfied (fair, transparent, accountable), the path to automation is clearer.*

**4. Stakeholder Value & Impact**

* **Who Benefits:** Identify key stakeholders (students, lecturers, admin staff, etc.). What benefit does AI offer each? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* **Potential Drawbacks:** What concerns might each stakeholder have? (Job impact, trust, quality drop?) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* **Communication:** Do stakeholders know about this AI plan and have input? Consider running a consultation or informing users. Any feedback so far? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* **Equity:** Will the AI service everyone fairly? (e.g. accessible to disabled users, works for international students) Any adjustments needed for inclusivity? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* **Value vs Effort:** Overall, does the AI use substantially improve something (speed, quality, access) for stakeholders? If the improvement is minor or unclear, is automation justified? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*If the AI doesn’t clearly add value or has significant stakeholder opposition, reconsider its use. Strong stakeholder benefits and acceptance support moving to AI-Assisted or Automated categories.*

**5. Human Element Requirement**

* **Empathy Needed:** Does the task involve emotional support, mentorship, or sensitive communication? If yes, humans should likely lead. Note the human touch points: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* **Judgment Needed:** Does the task require complex decision-making or moral judgment that humans excel at? Examples: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* **Tolerance for Automation:** Would a student or staff expect/accept this being automated? (e.g. maybe okay for scheduling, not okay for personal feedback) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* **Human Value Add:** Is there a benefit in keeping a human involved (like learning something by doing the task manually, or personalisation)? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* **Role of Human if AI is Used:** If AI is introduced, what *exactly* will humans still do? (e.g. validate results, handle exceptions, provide personal context) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*Use these answers to decide how much human oversight to retain. Tasks with high human element needs should be Human-Led or only lightly AI-assisted. Tasks with low need can be more automated, but ensure periodic human check-ins.*

**6. Feasibility & Implementation**

* **AI Solution Availability:** Is there an existing AI tool for this? \_\_\_\_ If yes, which? \_\_\_\_\_\_\_\_\_\_\_\_. If no, what would it take to develop one? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* **Integration:** Can it integrate with our systems (LMS, databases)? Do we need APIs or new infrastructure? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* **Cost:** What’s the cost (license or development and maintenance)? Budget approved? \_\_\_\_
* **Expertise:** Do we have the technical expertise to deploy and maintain it? Who will take ownership (department/role)? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* **Timeline:** How long to implement (including pilot and training)? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* **Training & Change Management:** Will staff/users need training? Plan for training: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_. How will we roll this out (pilot first, announcement)? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
* **Ongoing Maintenance:** Who will monitor performance and handle updates or issues long-term? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*If feasibility issues are significant (no tool available, very high cost, no expertise), you may need to delay or invest in preparation before using AI. If it’s readily feasible, proceed but ensure maintenance and monitoring responsibilities are clear.*

**Mapping to Decision Category:** After answering the above, determine which category fits best now:

* Mostly “human essential” and high-risk/ethics concerns → **Human-Led (Category 1)**
* Human needed, but AI can help in parts and is feasible → **AI-Assisted (Category 2)**
* Low risk, structured, AI proven effective, just needs oversight → **High Automation Potential (Category 3)**
* Very low risk, repetitive, AI is reliable and can run by itself → **Full Automation (Category 4)**

*(It’s okay if a task falls between two categories – choose the safer category to start, and you can always increase automation later.)*

**Final Check:** Before implementing, ensure to:

* Get necessary approvals (especially if high-risk AI under regulations).
* Inform and train stakeholders as needed.
* Set up monitoring and a feedback loop (schedule a review after X months).
* Document the decision (why this category, any conditions like “human must approve all AI outputs”).

By following this checklist, staff can systematically evaluate tasks and make informed decisions about integrating AI. The goal is to reap AI’s benefits where appropriate, while maintaining oversight and educational values. This ensures AI adoption at the university is done in a **responsible, transparent, and effective** manner ([7 principles on responsible AI use in education | World Economic Forum](https://www.weforum.org/stories/2024/01/ai-guidance-school-responsible-use-in-education/#:~:text=7,impact%20of%20AI)), in line with both institutional policy and global best practices.