AI Framework

Dated: 6/10/2025 **Ver: 0.2**

Introduction

As part of Columbia University's commitment to responsible innovation and technological excellence, the Enterprise Architecture (EA) team is collaborating closely with the Enterprise Applications team, the Emerging Technologies team, and the Risk Management team to establish a comprehensive framework for AI application development. This strategic initiative is grounded in the need to create a unified set of standards and best practices that ensure AI tools are effectively aligned with the university's broader technology landscape, institutional values, and operational goals. The core objective of this framework is to guide various stakeholders—developers, researchers, IT staff, and business units—in selecting, designing, and deploying AI solutions that are secure, scalable, ethically sound, and operationally sustainable (*Introducing Columbia AI*, 2023).

The framework emphasizes interoperability with Columbia's existing enterprise systems, such as the student information system, financial platforms, and HR solutions. To ensure maximum impact and usability, the framework incorporates input from both technical and non-technical teams, bridging academic innovation with operational requirements. It recognizes that AI has the potential to transform how Columbia delivers services, making them more intuitive, responsive, and personalized. Whether through intelligent chatbots, document summarization, or predictive analytics, AI solutions must be thoughtfully designed to enhance staff productivity while elevating the quality of student and faculty experiences.

A key tenet of the framework is governance — ensuring that all AI initiatives adhere to security, privacy, and compliance requirements from the outset. In partnership with Risk Management, the EA team is defining policies that mitigate risks such as data misuse, algorithmic bias, and model drift. The framework also promotes transparency and auditability, encouraging teams to document AI model choices, data sources, and expected outcomes. This ensures that AI solutions can be evaluated and trusted over time (*Generative AI Policy | Office of the Provost*, 2024).

In essence, this AI framework is not just a technical guide—it is a strategic compass that ensures Columbia University leverages artificial intelligence in a way that is intelligent, ethical, and institutionally aligned. It provides clarity amid rapid technological change and positions Columbia as a leader in purposeful AI adoption within higher education.

Strategy Pillars

- **Secure AI**: Ensures that all AI solutions are designed and deployed with strong safeguards for data privacy, cybersecurity, and compliance from the ground up.
- **Responsible AI**: Promotes ethical use of AI by emphasizing transparency, fairness, accountability, and alignment with institutional values.
- **Self-financing AI (Self-servicing AI)**: Focuses on building AI solutions that reduce manual workload, scale efficiently, and deliver measurable productivity gains that justify their investment.

AI Principles

- **AI First:** It is a guiding principle that encourages staff to proactively evaluate and integrate AI capabilities when designing new solutions or enhancing existing ones. It promotes a mindset where AI is considered early in the planning process—not as an afterthought—enabling smarter automation, improved user experiences, and data-driven decision-making. By embedding AI exploration into project lifecycles, this principle ensures Columbia stays ahead in leveraging innovative technologies that align with our institutional goals.
- **AI is Accountable**: AI systems must have clear ownership and auditability, with humans responsible for outcomes and oversight.
- **AI is Fair**: AI must treat all individuals equitably, minimizing bias and ensuring inclusive decision-making.
- **AI is Human-Centric**: AI should augment human capabilities, respect user needs, and prioritize user experience.
- **AI is Ethical**: AI development and use must align with institutional values, legal standards, and societal norms.
- **AI is Reliable**: AI systems should perform consistently, accurately, and as intended under expected conditions.
- **AI is Robust**: AI must be resilient to adversarial inputs, system failures, and changing environments.
- **AI is Scalable**: AI solutions should be designed to grow efficiently with institutional needs and data volumes.
- **AI is Adaptable**: AI should continuously learn and evolve to remain effective in dynamic academic and operational contexts.

Methodology

Adopt a structured approach to identifying, evaluating, and implementing AI opportunities aligned with institutional goals and ethical standards. Follow the steps below to ensure effective prioritization, scalable execution, and responsible adoption

- **Use Cases Prioritization**:

  Systematically evaluate AI opportunities by assessing their **business value**—such as impact on productivity, user experience, and alignment with strategic objectives—and **feasibility**, including data availability, technical complexity, and resource readiness. Leverage the prioritization matrix to identify high-potential initiatives and select use cases with **out-of-the-box capabilities** to reduce development time and accelerate delivery.

- **Decision Framework for Build Vs Buy**:

  Provide structured criteria to determine whether to develop AI solutions internally or adopt external tools.

- **Build Pilot Use Case for Scalability**:

  Start with a focused, low risk use case to validate value and scalability before full-scale deployment.

- **Use Composable Platform Architecture**:

  Leverage modular, interoperable systems that support flexible AI integration and rapid innovation. Links below provide guidance on use case realization, application of security framework, decision tree for AI tool selection, and Tool Evaluation

  - Use case realization - Provides Guidelines for use cases realization
  - Security Framework - Provides Framework and guideline for AI security
  - AI Decision Tree - Provide framework for Tool Selection
  - AI Toolset Evaluation - Provides framework for Tool Evaluation
  - Technology Stack Template - Captures AI footprint for the project in AI Section

- **Responsible AI at the Forefront**:

  Embed ethical, transparent, and accountable practices throughout the AI development lifecycle.

- **Vision and Principles for Responsible AI**:

  Anchor projects in clearly defined principles that reflect Columbia's values and societal responsibilities.

- **AI Literacy and Training**:

  Equip staff and stakeholders with the skills and understanding needed to effectively use and govern AI.

- **Enable Collaboration among Humans and Machines**:

  Design AI systems that complement human judgment and foster shared decision-making.

- **Apply Cost Management to AI**:

  Ensure AI investments are monitored, measured, and optimized for sustainable value delivery.

Architecture Blueprint

Projects and initiatives will continue to use the Solution Architecture templates to document their architectural blueprints. The Stack template from above section will enumerate AI tools used in the project. Additionally, teams may include optional artifacts—such as composable architecture models and decision trees—where applicable, to enhance their architectural documentation.

Appendix: Prioritization Categories (Business Value/Feasibility)

Use the prioritization categories below to identify use cases that align with out-of-the-box capabilities, minimizing development effort and accelerating delivery.

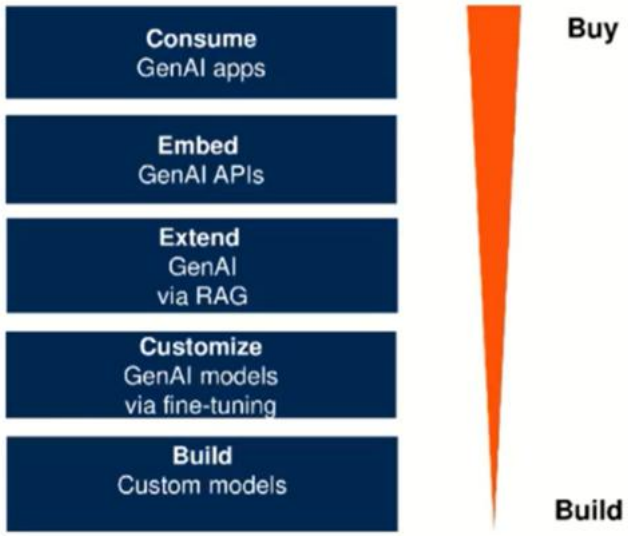
Priortization Categories (Business Value/Feasibility)

ID	Description	Data Domains	
1	Text Generation		☐
2	Image generation		☐
3	Code generation		☐
4	Summarization		☐
5	Audio Generation		☐
6	Contract Analysis		☐
7	Synthetic data		☐
8	Video Generation		☐
9	Custom Chatbot		☐
10	Search		☐
11	Talent Acquisition		☐
12	Model Training		☐
13	Custom Development		☐

Data Domains
SIS
HR
Financial
Academic
Research
Alumni
Facilities and Operations
Identity and Access Management
Risk and Compliance
Health and Wellness
Library and Acacdemic Resources
IT and Digital Infrastructure

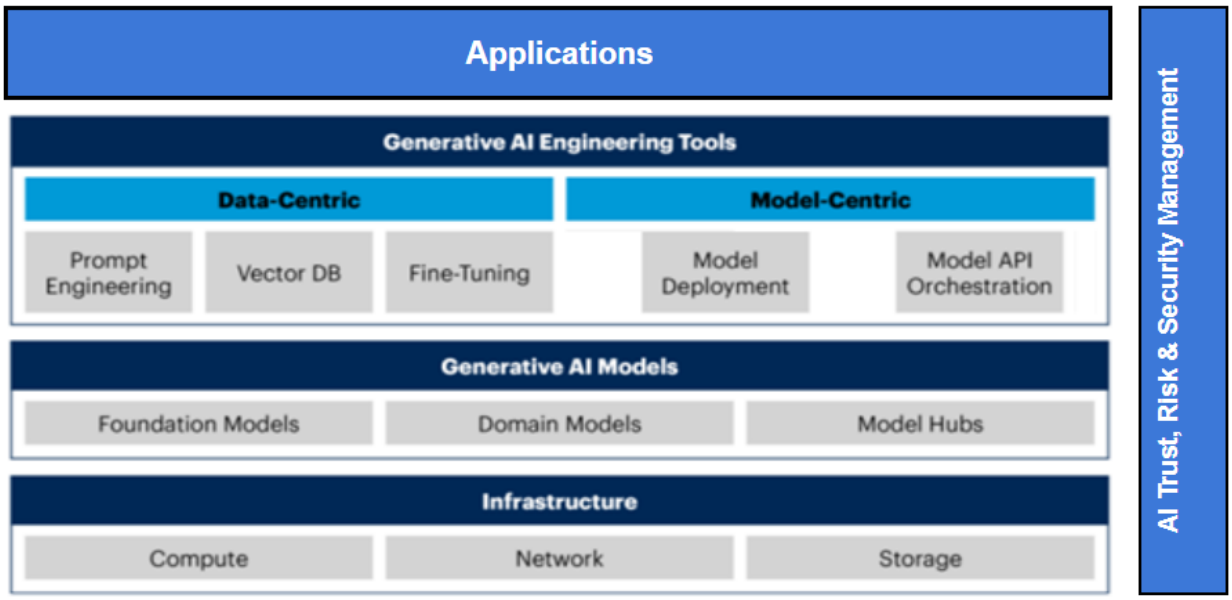
Appendix: Structured Criteria (Build vs Buy)

Use the structured criteria below to determine whether to adopt off-the-shelf components or develop a custom solution from scratch.



Appendix: Composable Platform Architecture

Leverage the composable architecture below to build solutions using AI components such as prompt engineering, vector databases, and foundational or domain-specific models.



Appendix: Use case realization

To effectively design and evaluate AI use cases, teams should align each initiative to Columbia's Generative AI Capability Schema. This framework outlines the core categories of capabilities and helps identify both technical and security considerations across each layer.

Capability Schema Overview

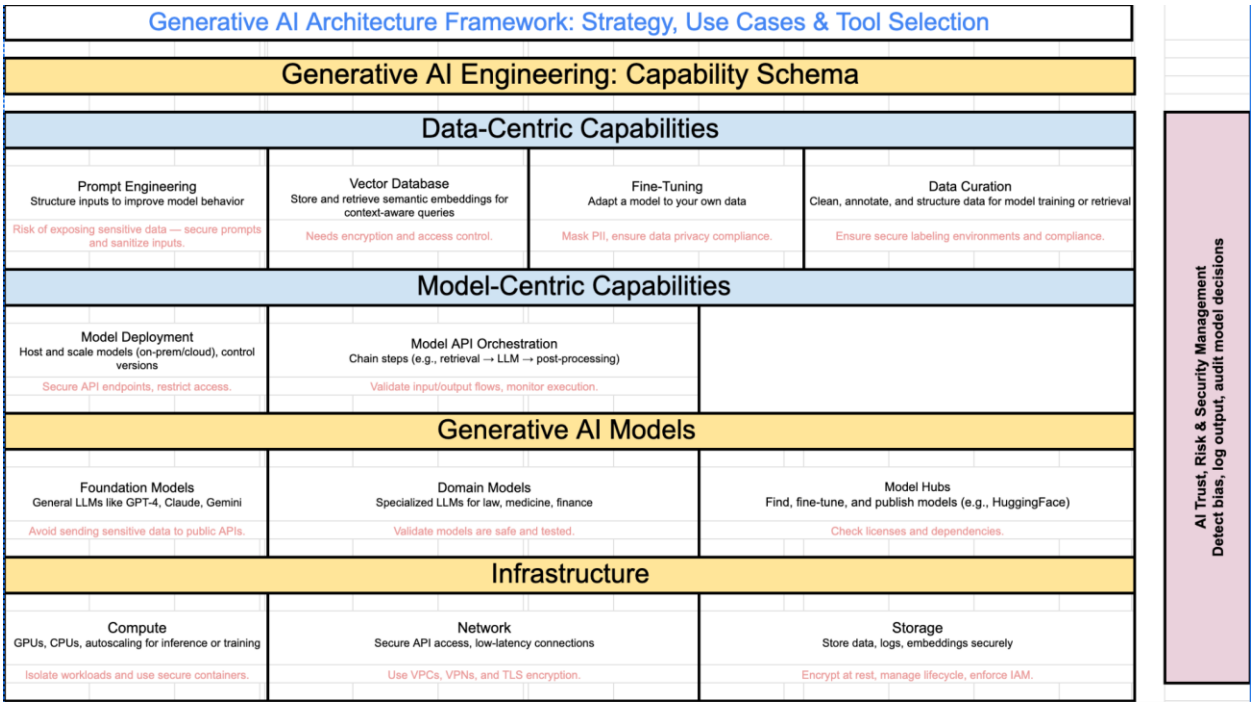


Figure 1: Generative AI Engineering Capability Schema. This diagram illustrates the layered approach to building AI systems: from data-centric engineering through model deployment and infrastructure, all surrounded by AI trust, risk, and security practices.

Security Considerations by Capability Category

Data-Centric Capabilities

- Prompt Engineering – Risk of exposing sensitive data; sanitize inputs.
- Vector Database – Requires encryption and access controls.
- Fine-Tuning – Must mask PII; comply with data privacy laws.
- Data Curation – Ensure labeling environments are secure and compliant.

Model-Centric Capabilities

- Model Deployment – Secure API endpoints and restrict access.
- Model API Orchestration – Validate inputs/outputs and monitor execution.

Generative AI Models

- Foundation Models – Avoid sending sensitive data to public APIs.

- Domain Models – Validate models for fairness and domain compliance.
- Model Hubs – Verify licensing and dependencies for compliance risks.

Infrastructure

- Compute – Use isolated, secure containers for workload execution.
- Network – Enforce VPC, VPN, and TLS for secure data in transit.
- Storage – Encrypt data at rest, enforce lifecycle and access controls.

AI Trust, Risk & Security Management

- Monitor model decisions, detect bias, log outputs, and ensure human oversight.

Use Case 1: Explain My Schedule

Element	Description
Problem Statement	Students need help understanding how their schedules are generated based on GPA, prerequisites, and program requirements.
Target Users	Students, Academic Advisors
AI Task Type	Natural Language Explanation (Summarization, Clarification)
Use Category	Consume, Embed
Model Type	Foundation Model
Capabilities Used	Prompt Engineering, Model API Orchestration, Secure Storage
Tool Options	GPT-4 (OpenAI/Azure), Claude (Anthropic), LangChain, AWS Bedrock
Security Notes	FERPA-sensitive inputs must be sanitized; outputs are encrypted and auditable
Business Impact	Improves student transparency, reduces advisor workload
Compliance Requirements	FERPA, access control, encryption at rest

Use Case 2: AI Advisor for Degree Planning

Element	Description
Problem Statement	Students need personalized course planning that considers academic history, GPA, and degree requirements.
Target Users	Students, Academic Advisors
AI Task Type	Planning, Recommendation, Explanation

Use Category	Embed, Extend
Model Type	Foundation Model, Domain Model
Capabilities Used	Data Curation, Vector Database, Prompt Engineering, API Orchestration, Secure Deployment
Tool Options	SOTA(State-of-the-art/foundational) Model from partners we have agreements/BAA with like GPT-4, Bedrock, Hugging Face, LlamaIndex, LangChain
Security Notes	Secure API endpoints, containerized model deployment, VPC network isolation
Business Impact	Enhances degree planning accuracy, supports underserved populations
Compliance Requirements	FERPA, secure access, data encryption, audit logging

Capability Mapping Summary

Capability Category	Capability	How It's Used
Data-Centric	Prompt Engineering	Design prompts that reflect user intent and academic rules
Data-Centric	Vector Database	Store & retrieve student profiles for similarity-based plan generation
Data-Centric	Data Curation	Clean, annotate academic records and course history
Model-Centric	Model Deployment	Run models in secure containers with access control
Model-Centric	API Orchestration	Chain logic: input → retrieval → LLM → output explanation
Infrastructure	Compute / Storage / Network	Use GPU-backed instances, encrypted storage (e.g., PostgreSQL, S3), and private APIs

Appendix: Security Framework

Risk Based AI Categorization

Columbia University categorizes AI technologies based on risk, impact, and regulatory obligations to ensure responsible deployment across academic, administrative, and clinical domains. Following a risk-based approach aligned with NIST AI RMF 1.0, higher-risk AI systems undergo stricter oversight and mitigation measures.

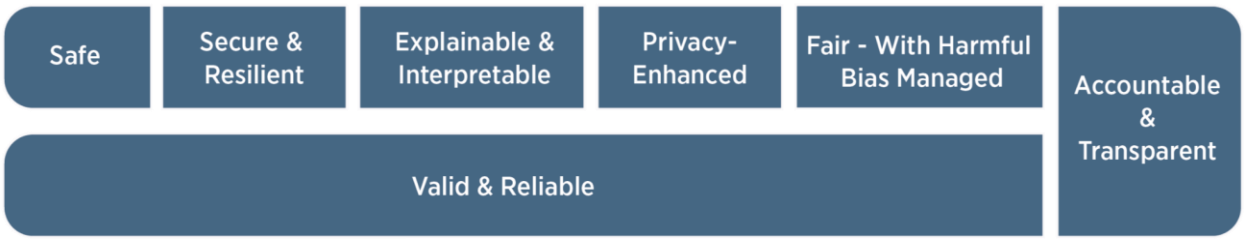
This classification ensures AI use remains ethical, secure, and compliant while supporting learning, administration, and clinical operations without compromising fairness, privacy, or institutional values.

AI Core Principles

Columbia University has established a structured framework to ensure the responsible, ethical, and transparent use of AI technologies across academic, administrative, and clinical domains. These principles serve as the foundation for AI adoption, ensuring that all AI-related decisions align with the university's core values, regulatory obligations, and commitment to human well-being.

A core tenet of this governance framework is Ethical AI Use, emphasizing the responsible development, deployment, and oversight of AI technologies. Guided by the **NIST AI Risk Management Framework (RMF) 1.0**, these principles ensure AI systems operate safely, securely, reliably, fairly, transparently, and with accountability while enhancing privacy. (*Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, 2023)

These core principles ensure AI-related decisions align with the university's mission, comply with regulations, and minimize potential risks and unintended consequences.



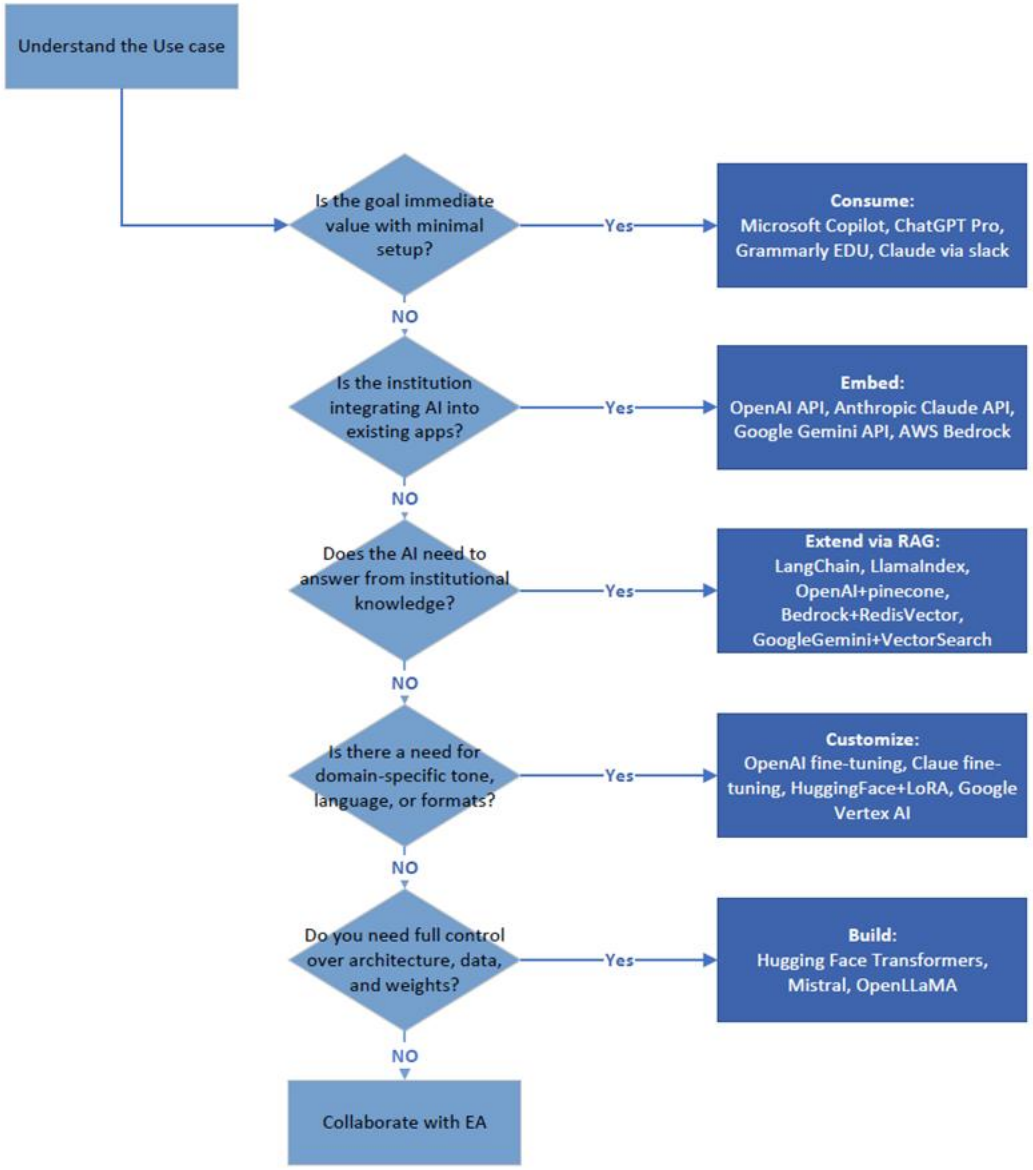
NIST AI RMF 1.0 - characteristics of trustworthy AI

Core Principles	Description	Example
Valid and Reliable	AI systems must produce accurate and consistent results across various scenarios. They should be thoroughly tested and validated to ensure reliability and minimize biases.	An AI model used for loan approvals becomes unreliable over time due to changing applicant data, leading to incorrect decisions and financial losses.
Safe	The use of AI must prioritize safety for users, the institution, and the broader community. AI systems must be designed and implemented with robust safety mechanisms to minimize risks and potential harm.	An autonomous vehicle experiences a system failure and causes a collision
Secure and Resilient	AI systems must be protected against cybersecurity threats, unauthorized access, and data breaches. Robust security measures, including data encryption and access controls, must be implemented to ensure the confidentiality, integrity, and availability of AI systems and data.	Facial recognition systems storing sensitive data without adequate safeguards.
Accountable and Transparent	Clear ownership and accountability for AI systems are essential. AI accountability requires clear identification of roles and responsibilities for everyone involved in developing, deploying, and using AI systems. Decision-making processes should be transparent and auditable. Stakeholders should have access to information about the AI systems used and their intended purposes.	An innocent person is convicted of a crime based on evidence of AI driven decision.
Explainable and Interpretable	AI systems and their outputs should be understandable and interpretable, enabling users to understand how decisions are made. This promotes trust and facilitates responsible use.	Loan approval systems providing no clear explanation for denials, undermining trust
Privacy-Enhanced	AI systems must be designed and	AI tools processing PII data

	implemented with privacy at the forefront. Data collection, use, and storage should comply with applicable privacy laws and regulations, such as FERPA and GDPR.	without adhering to privacy laws like GDPR, HIPAA, etc
Fair - Mitigating Bias	AI systems must be developed and deployed in a way that avoids bias and ensures equitable outcomes for all users. Continuous monitoring and evaluation are crucial to identify and mitigate potential biases.	A hiring algorithm favoring certain demographics due to historical biases in recruitment data with which it was trained.

Appendix: AI Decision Tree

Use the decision tree below to identify the set of AI tools best suited for implementing your use case.



Appendix: AI Toolset Evaluation

To ensure objective and consistent selection of AI tools, Columbia proposes a scoring framework based on six key criteria. Each tool is scored on a scale of 1 to 10 using standardized rubrics, and then weighted to reflect institutional priorities.

Evaluation Criteria

Criterion	What It Measures
Performance	Speed, scalability, and accuracy under real-world workloads
Cost	Licensing, usage, infrastructure cost, and transparency
Customization	Ability to fine-tune or configure for domain-specific use
Compliance & Security	Support for regulatory/security standards (FERPA, HIPAA, etc.)
Integration & Tooling	Compatibility with systems, APIs, SDKs, and documentation
Community & Support	Docs, vendor responsiveness, and community activity

Criterion	Weight
Performance	25%
Cost	20%
Customization	15%
Compliance & Security	15%
Integration & Tooling	15%
Community & Support	10%

Default Evaluation WeightsWeighted Score Formula

Each criterion is multiplied by its assigned weight and summed to produce the final score:

Final Score = (Performance × 0.25) + (Cost × 0.20) + (Customization × 0.15) + (Compliance × 0.15) + (Integration × 0.15) + (Support × 0.10)

Scoring Rubrics by Criterion

Performance

Score	Description
10	Real-time, high-throughput, state-of-the-art accuracy
8–9	Fast and reliable under normal load, consistently accurate
6–7	Adequate for most tasks but slower or slightly less accurate
4–5	Noticeable delay or inconsistent accuracy
1–3	Slow, unreliable, or error-prone under load

Cost

Score	Description
10	Free and open-source
9	Minimal cost (e.g., <$10/month)
7–8	Predictable pay-as-you-go or modest enterprise plan
5–6	Expensive but justifiable for business value
1–4	High or unpredictable costs, opaque pricing

Customization

Score	Description
10	Full training and fine-tuning support with custom architectures
8–9	Fine-tuning supported or modular adapter integration
6–7	Prompt engineering and plug-ins only
4–5	Minor tweaking via API settings
1–3	Closed system, no customization

Compliance & Security (preferable with BAA agreements)

Score	Description
10	Certified for FERPA, HIPAA, GDPR, SOC 2
8–9	Meets major compliance standards (e.g., GDPR, SOC 2)
6–7	Some security, partial certifications
4–5	Basic auth, no formal compliance
1–3	Unknown or insecure

Integration & Tooling

Score	Description
10	SDKs, plug-ins, CLI tools, multi-language support
8–9	Well-documented APIs, easy integration
6–7	APIs exist but require effort or workarounds
4–5	Minimal tooling or poorly documented
1–3	Manual integration required

Community & Support

Score	Description
10	Strong open-source community, active updates, SLAs
8–9	Good docs, quick support response
6–7	Some tutorials or community presence
4–5	Sparse docs or outdated libraries

1–3	Dead project, no help available

Appendix: Technology Stack Template

The following section is added to Technology Stack Template

AI Section		
Foundational Model		
Prompt Engineering		
Vector DB		
Fine Tuning		

Appendix: Solution Architecture Template

Use the CUIT-approved Solution Architecture Template to document the blueprint for AI integration within the project.

References

- *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. (2023, January 1). NIST Technical Series Publications. Retrieved May 27, 2025, from https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf
- *Generative AI Policy | Office of the Provost*. (2024, Nov 13). Office of the Provost. Retrieved May 27, 2025, from https://provost.columbia.edu/content/office-senior-vice-provost/ai-policy
- *Introducing Columbia AI*. (2023, Nov 13). Columbia University. (2023). https://news.columbia.edu/news/introducing-columbia-ai?utm_source=MarketingCloud&utm_medium=email&utm_campaign=20241118_Spotlight_AI_Faculty_Staff