A DECISION MODEL FOR SAFE INSTITUTIONAL DATA USE WITH GENERATIVE AI IN HIGHER EDUCATION

Alexandru DICEA[1], Ciprian DOBRE[2]

Abstract. *The rapid adoption of generative artificial intelligence (GenAI) in higher education creates a practical data-governance problem at the point of use. This paper proposes a compact, policy-grounded decision model combining institutional data characteristics, GenAI service assurance and approved scope, and intended use. Before disclosure, the model returns one of six next actions: proceed, transform and reassess, use an approved service, seek institutional review, withhold submission, or verify missing information. Analytical evaluation using 20 constructed higher-education scenarios indicates internally coherent behaviour and preserves the boundaries among uncertainty, incompatibility, escalation, and prohibition under the reference policy profile. The model is intended as a user-facing decision layer beneath institutional AI governance and does not replace existing privacy, cybersecurity, governance, or compliance mechanisms.*

Keywords: generative AI; higher education; institutional data governance; Shadow AI; decision support

DOI [10.56082/annalsarsciinfo.2026.2.5](https://doi.org/10.56082/annalsarsciinfo.2026.2.5)

1. Introduction

Generative artificial intelligence (GenAI) is increasingly used in higher education for drafting, summarization, translation, analysis, teaching support, research, and administrative work [1]–[3]. Its accessibility allows many of these activities to be performed through externally hosted services with little or no institutional infrastructure. At the same time, the information supplied to such services may be internal, confidential, personally identifiable, research-sensitive, contractually protected, or otherwise governed by institutional rules. The practical governance problem therefore arises at the point where a user who already legitimately handles institutional information is about to submit it to a GenAI service.

Universities increasingly respond through GenAI policies, approved-service lists, privacy and security guidance, research rules, and broader governance frameworks [1], [4]–[6]. These mechanisms are necessary, but they do not eliminate the user-facing interpretation problem. A concrete decision may depend simultaneously on the data being submitted, the assurance and approved scope of the selected service, the purpose of use, and other institutional constraints. An approved service may not

[1]Eng. PhD candidate, Faculty of Automatic Control and Computers, POLITEHNICA Bucharest, Romania, e-mail: alex.dicea@upb.ro
[2]Prof. Dr. Eng. Habil., Faculty of Automatic Control and Computers, POLITEHNICA Bucharest, Romania, e-mail: ciprian.dobre@upb.ro

be approved for every data class or activity, while a task that appears routine may require additional review if it contributes to a consequential institutional decision.

Existing work addresses several adjacent layers. Shadow IT and Shadow AI research examines unauthorized technology use and organizational governance [7]–[9], [12]–[16]; higher-education research analyses and develops institutional GenAI policies [1], [4]–[6]; privacy and security studies examine data-disclosure concerns [6], [17]; and policy-aware systems support organizational data-access decisions [11]. Cherner et al. [10] additionally provide an end-user decision tree for evaluating GenAI tools in higher education. These contributions establish important upstream governance mechanisms, but they address less directly the operational decision immediately before a specific disclosure.

The problem considered here is therefore narrower than general AI governance and narrower than Shadow AI detection or remediation. It can be expressed as a practical question:

> *Can this institutional information be submitted to this GenAI service for this purpose, and what is the next action before disclosure?*

The study formalizes that question through the following research question:

> *How can a higher-education user determine the appropriate next action, before submitting institutional information to a generative-AI service, for a specific combination of data characteristics, service assurance, and intended use under applicable institutional constraints?*

The question deliberately seeks a next action, not a binary allowed/not-allowed verdict. This distinction is important because many realistic cases require an intermediate step: the user may need to transform and reassess information, move to an approved service, seek institutional review, or verify missing information rather than simply proceed or stop.

The conceptual form of the proposed model is:

$$D \times S \times P \xrightarrow{\Pi} A \tag{1}$$

where D represents data characteristics, S service assurance and approved scope, P intended use, Π applicable institutional policy and constraints, and A the next action before disclosure. The more explicit form introduces contextual facts C:

$$F_{\Pi}(D, S, P, C) \rightarrow A \tag{2}$$

The contribution is deliberately narrow. It does not claim novelty from providing a generic GenAI decision aid for higher education, from connecting Shadow AI to privacy risk, or from introducing policy-aware AI decision support in general. Instead, it concerns the policy-grounded mapping of a concrete institutional-data disclosure configuration to an actionable next step before submission. The model is intended to operationalize institutional governance at the user-facing disclosure

point, not to replace the governance, privacy, cybersecurity, legal, or compliance mechanisms from which its rules are derived.

2. Background and Related Work

2.1 Shadow AI, higher-education governance, and data disclosure

Shadow AI extends the established Shadow IT problem to informal or unauthorized organizational use of AI. Shadow IT research has examined drivers, consequences, and governance challenges of technology adopted outside formal IT control [12]–[15], while recent Shadow AI work adds GenAI-specific concerns such as third-party processing, data disclosure, opaque processing chains, probabilistic behaviour, and AI-generated influence on decisions [7]–[9], [16].

Within higher education, Zentner and West [8] frame unauthorized employee AI use as a socio-technical governance problem. Sebastian [9], through the Digital Shadow AI Risk Theory (DART), focuses more directly on data disclosure, privacy, and organizational Shadow AI risk. DART nevertheless remains an organizational risk-oriented framework rather than a point-of-use mechanism that maps a specific data-service-purpose configuration to the next pre-disclosure action. Ojowa et al. [16] propose an operational framework centred on detection, risk classification, governance, and remediation. These approaches are directly relevant to the context of the present study, but their primary decision locus remains organizational: identifying or governing unauthorized AI use rather than determining the immediate user action before a specific institutional-data disclosure.

Higher-education institutions also increasingly develop GenAI policies covering teaching, research, assessment, administration, privacy, and security [1], [4]–[6]. Li et al. [4] derive a university policy-development framework from cross-national analysis, while Nguyen [5] proposes a stepwise method for operationalizing AI governance in higher education. Ng et al. [6] examine how university GenAI guidance addresses privacy and security. This work establishes that institutional guidance is no longer absent; rather, a significant challenge is translating broad policy and safeguards into a concrete decision at the point of use.

Institutional classification and service approval may operate at different levels, with approval limited to particular data classes, purposes, or configurations. Point-of-use decisions therefore require coordination across governance layers rather than a single universal rule.

2.2 Closest decision-oriented approaches

Two approaches are particularly important for delimiting the contribution.

Cherner, Foulger, and Donnelly [10] provide an end-user decision tree for ethical assessment of GenAI tool adoption in higher education. The present study begins later: the user already legitimately holds institutional information and must

determine the next action before disclosing it to a particular service for a particular purpose.

The distinction can be summarized as: Cherner et al.: purpose + tool + ethical considerations → tool adoption/evaluation, whereas the present study addresses institutional data + service + purpose + Π → next pre-submission action.

The second important comparator is Al Mandalawi et al. [11]. Their policy-aware GenAI approach interprets natural-language data-access requests against organizational policies and metadata and returns auditable access decisions. This establishes that policy-aware AI decision mechanisms already exist. The protected operation is nevertheless different. Their work concerns authorization to access organizational data; the present study begins after legitimate access has already been established and governs whether already-held institutional information should be disclosed to GenAI.

The difference is therefore: Al Mandalawi et al.: request + policy + metadata → data-access decision, versus present study: institutional data + service + purpose + Π → next disclosure action.

A compact comparison is shown in Table 1.

Table 1. Positioning relative to the closest prior-work families.

Prior-work family	*Primary decision locus*	*Difference from this study*
Shadow AI governance [7]–[9], [16]	Organizational detection, risk, governance, remediation	Not a point-of-use pre-disclosure decision
HE GenAI governance [1], [4], [5]	Institutional policy development and operationalization	Develops governance rather than applying it to a specific disclosure
University privacy/security guidance [6]	Content and implementation of institutional guidance	Analyses guidance rather than returning a next action
Cherner et al. [10]	End-user ethical assessment of GenAI tool adoption	Tool-adoption decision rather than policy-grounded institutional-data disclosure
Al Mandalawi et al. [11]	Policy-aware enterprise data-access authorization	Access authorization rather than disclosure of already-held data to GenAI

The residual gap is therefore not the absence of institutional guidance, but its operational interpretation for a specific intended disclosure. The present contribution occupies this downstream point in the governance chain and treats policy as an active part of the decision rather than as general context.

3. Method

3.1 Research design and requirement derivation

The study followed a design-oriented analytical approach comprising five linked stages: focused source synthesis, requirement derivation, model construction, scenario-based analytical evaluation, and boundary/failure analysis. It involved no human participants, institutional deployment, or statistical evaluation; accordingly, evaluation concerns internal analytical consistency and boundary behaviour rather than empirical effectiveness.

The focused source base combined foundational and recent work on Shadow IT and Shadow AI [7]–[9], [12]–[16], higher-education GenAI governance and privacy/security guidance [1], [4]–[6], policy-aware decision approaches [11], relevant standards and guidance [17]–[19], and institutional GenAI practice [20]–[23]. The synthesis was intentionally focused rather than systematic; its purpose was to identify what a pre-submission decision model must represent rather than to estimate prevalence or effect sizes.

Candidate requirements were examined according to whether they materially changed the point-of-use decision. A concept was retained as a principal dimension only when it could not be represented adequately as an attribute or policy constraint. This procedure produced three principal dimensions: data characteristics D, GenAI service assurance and approved scope S, and intended use P. Institutional rules were represented separately by the policy layer Π, while contextual facts needed by particular rules were represented by C.

This separation keeps the user-facing structure compact while preserving institution-specific classification, service-assurance, authorization, and review rules.

3.2 Reference policy profile and scenario set

Because the model is policy-grounded rather than universal, analytical evaluation required an explicit reference policy profile Π_{ref}. It was constructed from handling principles identified in the source synthesis and institutional-guidance review [17], [18], [20]–[23]. The profile included reference rules for public, internal, confidential, and restricted information; approved and unapproved services; data minimization and de-identification; unpublished research and intellectual property; consequential institutional decision support; uncertainty; and explicit institutional prohibitions.

The profile is deliberately not presented as a universal university policy or legal-compliance model. In practice, an institution would instantiate Π from its own information-classification, privacy, cybersecurity, research, procurement, contractual, legal, and AI-governance rules. Standards and regulation can therefore inform institutional governance, but they do not map directly to ALLOW or DO_NOT_SUBMIT. The EU AI Act, for example, is relevant regulatory context [24], not a direct action engine for this model.

A set of 20 constructed higher-education scenarios was then developed for analytical coverage rather than statistical representativeness. It spans teaching and learning, administration, admissions, assessment, research, student information, unpublished research and intellectual property, and institutional legal material. The scenario set contains ordinary cases, uncertainty, escalation, explicit prohibitions, and borderline configurations.

Controlled contrasts were included in which one decision-relevant component changes while the remaining case remains substantially constant. These contrasts examine whether changes in D, S, or P produce coherent changes in A. For every scenario, the expected action was derived by applying Π_{ref} to the documented conditions. These expected actions are analytical reference outcomes, not empirical ground truth.

3.3 Evaluation and boundary analysis

Each scenario was evaluated using the formal decision sequence. The procedure assessed whether the information required by policy was known, whether the use was explicitly prohibited, whether the selected service was compatible with the data and purpose, whether legitimate transformation could resolve an incompatibility, and whether institutional review was required.

The resulting A_{model} was compared with $A_{expected}$. Evaluation focused on three properties:

1. Action coverage — whether realistic constructed cases exercise all six action classes.

2. Controlled sensitivity — whether isolated changes in D, S, or P produce coherent action changes.

3. Boundary behaviour — whether the model distinguishes missing information, known incompatibility, review requirements, and explicit prohibition.

Agreement between A_{model} and $A_{expected}$ is interpreted as analytical consistency, not predictive accuracy or independent validation. This qualification is important because the scenarios and their expected actions were derived using the same reference-policy logic that informed the model. The evaluation therefore asks whether the formal procedure behaves consistently with its stated policy-grounded requirements, not whether it predicts an independently observed empirical outcome.

A separate boundary analysis distinguished failures of the formal decision logic from failures caused by inaccurate inputs or inadequate governance information. It then examined failure modes that may remain even when the formal logic is internally correct: data misclassification, omitted protection tags, incomplete de-identification, purpose drift, stale service assessments, unclear approved scope, policy gaps or conflicts, incomplete contextual information, cumulative disclosure, and interpretation of ALLOW. This analysis produced two model hardenings. First, policy silence is treated as uncertainty rather than permission. Second, decision validity is configuration-specific:

$$\Delta(D, S, P, C, \Pi) \Rightarrow \text{REASSESS.} \tag{3}$$

Thus, a previous decision cannot be reused after a material change in the data, service, purpose, context, or applicable policy.

4. Proposed Decision Model

4.1 Rationale and architecture

The model addresses a single point-of-use question: given institutional information that a higher-education user already legitimately handles, a selected GenAI service, and an intended use, what should the user do before submitting the information? The model therefore evaluates disclosure to GenAI, not authorization to access the underlying institutional data.

It is intentionally action-oriented rather than score-oriented. A numerical risk score may indicate relative concern but does not necessarily tell a user what to do next. The present model instead translates applicable institutional rules into one of six user-facing actions.

In this study, safe means acceptable under applicable institutional rules, known service-assurance conditions, identified constraints, and the configuration assessed at the time of decision. It does not imply zero residual risk or guaranteed legal compliance.

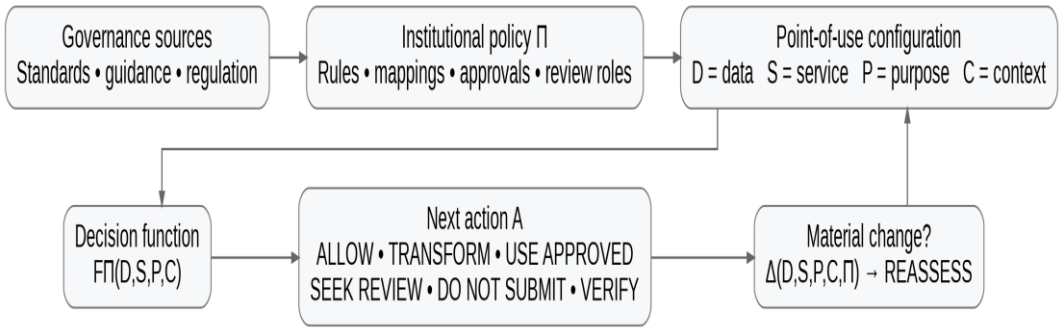

Fig. 1. Point-of-use decision architecture.

Governance sources are interpreted into the institution-specific policy layer Π. The evaluated configuration combines D, S, P, and contextual facts C; the decision function returns A; and any material change triggers reassessment.

4.2 Model components

The model components are summarized in Table 2.

Table 2. Components of the decision model.

Component	*Meaning*	*Principal representation*
D	Institutional data profile	Class: PUBLIC, INTERNAL, CONFIDENTIAL, RESTRICTED; tags: PERSONAL_DATA, STUDENT_RECORD, EMPLOYEE_DATA, RESEARCH, UNPUBLISHED_RESEARCH, INTELLECTUAL_PROPERTY, CONTRACTUAL, PRIVILEGED, OTHER_PROTECTED
S	GenAI service assurance and approved scope	S0 UNASSESSED; S1 EXTERNAL/UNAPPROVED; S2 INSTITUTION-APPROVED EXTERNAL; S3 INSTITUTION-CONTROLLED; approved scope and assurance properties
P	Intended-use profile	Processing operation; institutional use; consequentiality
Π	Applicable institutional rules	Prohibitions; compatibility; transformation; review; policy coverage
C	Contextual facts	Authorization; role; existing approval; data ownership; other rule-specific facts
A	Next pre-submission action	ALLOW, TRANSFORM, USE_APPROVED, SEEK_REVIEW, DO_NOT_SUBMIT, VERIFY

The internal representations remain compact:

$$D = (d_c, d_t), \qquad S = (s_c, s_a, s_p), \qquad P = (p_o, p_u, p_q). \tag{4}$$

Here, d_c is the normalized data class and d_t the protection-tag set; s_c is the service-assurance class, s_a the approved scope, and s_p the relevant assurance properties. For purpose, p_o uses the compact operation families TRANSFORM, GENERATE, ANALYZE, TECHNICAL, and OTHER; p_u distinguishes TEACHING_LEARNING, RESEARCH, ADMINISTRATION, COMMUNICATION_PUBLICATION, ASSESSMENT_DECISION, and OTHER; and:

$$p_q \in \{\text{ORDINARY},\ \text{CONSEQUENTIAL}\} \tag{5}$$

The data classes are normalized rather than universal. Institutions can retain their own classification schemes and map local categories to the normalized decision vocabulary. Protection tags describe why information may require additional handling even within the same general classification.

The service profile is similarly not a linear safety score. An S2 or S3 service is not automatically acceptable for every class of data or every purpose. Relevant service properties may include retention, provider reuse or training, third-party processing, contractual assurance, institutional administrative control, and processing boundary [17], [20], [21], [23]. The decisive question is whether the service's approved scope is compatible with the particular D and P under Π.

The purpose profile combines the operation being performed, the institutional use, and consequentiality. A use is consequential when GenAI contributes substantively to decisions such as admission, grading, employment, disciplinary action, or comparable decisions affecting individuals. Consequentiality remains part of P because the same data and service may legitimately produce different actions when the intended use changes.

4.3 Policy predicates and action semantics

The policy layer evaluates at least:

$$\begin{aligned} &\text{known}_{\Pi}(D,S,P,C) \\ &\text{covered}_{\Pi}(D,S,P,C) \\ &\text{prohibited}_{\Pi}(D,S,P,C) \\ &\text{compatible}_{\Pi}(D,S,P,C) \\ &\text{transformable}_{\Pi}(D,P,C) \\ &\text{reviewRequired}_{\Pi}(D,S,P,C) \end{aligned} \tag{6}$$

known_Π asks whether the decision-relevant facts are sufficiently known. *covered_Π* asks whether institutional policy sufficiently covers the case. *prohibited_Π* captures hard institutional restrictions. *compatible_Π* determines whether the selected service can be used with the relevant data and purpose. *transformable_Π* asks whether a legitimate transformation can remedy incompatibility without defeating the purpose. *reviewRequired_Π* identifies cases requiring institutional escalation.

The six actions have distinct meanings:

• ALLOW — no identified applicable institutional constraint prevents submission under the assessed configuration; this is not a zero-risk or legal-compliance guarantee.

• TRANSFORM — apply a legitimate transformation such as minimization, redaction, or de-identification, then reassess.

• USE_APPROVED — replace the current service with an institution-approved alternative whose scope is compatible with the data and purpose.

• SEEK_REVIEW — obtain the required institutional approval or specialist review before proceeding.

• DO_NOT_SUBMIT — an applicable hard prohibition prevents the evaluated disclosure.

• VERIFY — required information is missing, uncertain, stale, contradictory, or insufficiently covered by policy.

These actions are not severity levels. VERIFY, for example, indicates uncertainty rather than a known incompatibility, while DO_NOT_SUBMIT denotes an explicit prohibition.

4.4 Decision logic, reassessment, and example

The hardened function is:

In the decision function below, the symbol ¬ denotes logical negation ("not"). For example, $\neg\,\text{compatible}_{\Pi}(D,S,P,C)$ means that the selected service is not compatible with the assessed data-purpose configuration under the applicable institutional policy.

$$F_{\Pi}(D,S,P,C) = \begin{cases} \text{VERIFY}, & \neg\,\text{known}_{\Pi} \\ \text{VERIFY}, & \neg\,\text{covered}_{\Pi} \\ \text{DO NOT SUBMIT}, & \text{prohibited}_{\Pi} \\ \text{TRANSFORM}, & \neg\,\text{compatible}_{\Pi} \wedge \text{transformable}_{\Pi} \\ \text{USE APPROVED}, & \neg\,\text{compatible}_{\Pi} \wedge \neg\,\text{transformable}_{\Pi} \\ \text{SEEK REVIEW}, & \text{reviewRequired}_{\Pi} \\ \text{ALLOW}, & \text{otherwise} \end{cases} \tag{7}$$

Where service compatibility itself cannot be established:

$$\text{compatible}_{\Pi}(D,S,P,C) = UNKNOWN \rightarrow VERIFY \tag{8}$$

The ordering is procedural: uncertainty is resolved first, explicit prohibition dominates when known, incompatibility triggers remediation, review requirements trigger escalation, and ALLOW is reached only when no identified rule blocks the configuration.

TRANSFORM, USE_APPROVED, SEEK_REVIEW, and VERIFY are transitional actions; after the required change, the case must be reevaluated. A practical implementation can return *R=(A,E)*, where *E* records the reason, applicable rule, and, where relevant, review authority.

For example, internal administrative material submitted to an unapproved external service yields USE_APPROVED when policy requires an approved service:

$$S1 \rightarrow \text{USE APPROVED.} \tag{9}$$

After selecting an S2 service with compatible approved scope, the case is reassessed; service substitution is not final authorization. If no additional prohibition or review requirement applies:

$$S2 \rightarrow \text{REASSESS} \rightarrow \text{ALLOW.} \tag{10}$$

The example illustrates that the model returns the next policy-consistent action rather than a permanent safe/unsafe classification.

5. Scenario-Based Evaluation

5.1 Aggregate results

All 20 constructed scenarios produced the analytically derived reference action. This result is interpreted as internal analytical consistency under Π_{ref}, not as predictive accuracy or independent empirical validation.

All six action classes were exercised: ALLOW in 6 scenarios, TRANSFORM in 2, USE_APPROVED in 4, SEEK_REVIEW in 3, VERIFY in 3, and DO_NOT_SUBMIT in 2. Twelve of the twenty scenarios therefore produced an intermediate or remedial action rather than a terminal allow/prohibit result. This distribution is a property of the deliberately constructed scenario set, not an estimate of how frequently such outcomes occur in real higher-education use.

5.2 Controlled contrasts

Three controlled contrasts illustrate how the model reacts to changes in its principal dimensions.

Data variation. SC04 yields TRANSFORM for identifiable student material submitted to a service approved only for de-identified teaching material; after legitimate de-identification and reassessment, SC05 yields ALLOW:

$$\Delta D \Rightarrow \Delta A, \tag{11}$$

Transformation therefore requires reassessment rather than implying permission.

Service variation. SC02→SC03 changes from USE_APPROVED to ALLOW when an unapproved external service is replaced by one with compatible institutional approval; SC10→SC11 shows the same boundary for unpublished research/IP. Thus, institutional assurance and approved scope, not product identity alone, affect the action.

Purpose variation. With the same confidential admissions data and compatible service, changing the purpose from ordinary administration (SC08) to consequential admissions decision support (SC09) changes ALLOW to SEEK_REVIEW:

$$\Delta P \Rightarrow \Delta A. \tag{12}$$

Purpose and consequentiality therefore remain independent decision dimensions.

5.3 Boundary cases and analytical interpretation

Table 3 summarizes the broader evidence.

Table 3. Compact analytical evidence.

Evidence type	*Scenario(s)*	*Controlled or boundary condition*	*Action / transition*	*Interpretation*
Data variation	SC04 → SC05	Identifiable → de-identified	TRANSFORM → ALLOW	Transformation requires reassessment
Service variation	SC02 → SC03	Unapproved → approved compatible service	USE_APPROVED → ALLOW	Service scope affects action
Research/service	SC10 → SC11	Unapproved → approved for research/IP	USE_APPROVED → ALLOW	Service scope affects action
Purpose variation	SC08 → SC09	Ordinary admin. → consequential admissions	ALLOW → SEEK_REVIEW	Purpose affects action
Approved-scope boundary	SC06 → SC07	Incompatible → compatible approved scope	USE_APPROVED → ALLOW	Approval alone is insufficient
Uncertainty	SC12–SC14	Unknown service, data class, or scope	VERIFY	UNKNOWN is not permission
Strong assurance + review	SC17–SC18	Institution-controlled, but review required	SEEK_REVIEW	S3 does not imply ALLOW
Transformability	SC15 → SC16	Identity unnecessary → required	TRANSFORM → USE_APPROVED	Depends on purpose/data
Explicit prohibition	SC19–SC20	Hard institutional prohibition	DO_NOT_SUBMIT	Distinct boundary state

Boundary cases preserve four distinct states: missing or stale information → VERIFY; known incompatibility → TRANSFORM or USE_APPROVED; required authorization → SEEK_REVIEW; and explicit prohibition → DO_NOT_SUBMIT. These states require different user actions.

The evaluation therefore supports only three claims: all six actions are internally representable; the formal logic responds coherently to controlled changes in D, S, and P; and important boundaries are preserved rather than collapsed into a single risk score. Boundary analysis also shows that a formally consistent result can still be misleading when inputs or governance information is wrong, stale, or incomplete. For this reason, policy silence is uncertainty rather than permission, and decision validity is configuration specific.

6. Discussion

6.1 Interpretation and user implications

The model's practical value lies in distinguishing why a submission may not yet proceed rather than returning only a binary judgment. The scenarios also show that service approval or institutional control alone is insufficient; approved scope and intended use remain necessary.

From the user perspective, the model condenses several policy questions into a structured sequence:

What data? + Which service? + For what purpose? → What should I do next?

This structure may reduce interpretive burden, although usability is not empirically tested here. After any transitional action, the modified configuration must be reassessed.

ALLOW remains narrow: it means only that no identified applicable institutional rule prevents submission for the assessed configuration, not that the activity is risk-free, legally compliant across jurisdictions, permanently acceptable, or immune to changes in service terms or policy.

6.2 Institutional operationalization

The model shifts part of the operational burden from individual users to institutional governance. Users should not need to inspect vendor contracts, retention terms, training or reuse practices, subprocessors, or technical controls for every prompt. Institutions can assess these properties centrally and expose them through a maintained GenAI Service Registry or equivalent mechanism [20], [21], [23].

Such a registry could record:

• service assurance class;

• approved data classes;

• approved purposes;

• relevant contractual or technical restrictions;

• assessment date;

• validity or reassessment date.

Institutional data-classification schemes can likewise be mapped centrally to the normalized D vocabulary. This produces a clear division of responsibility: the institution defines policy, mappings, service assurance, approved scope, and review

authorities; the user identifies the data, service, and intended purpose; the model determines the next action; and the competent specialist resolves cases requiring review.

The reusable structure is therefore $F_{\Pi}(D, S, P, C)$, not a universal policy. Different institutions may legitimately produce different actions for the same nominal configuration because Π, contractual arrangements, legal interpretations, approval processes, and local precedence rules differ. The model should thus be understood as decision support beneath governance, not as autonomous policy enforcement.

6.3 Relationship to Shadow AI and transferability

Shadow AI is an important context, but the model is not a complete Shadow AI control framework. Whereas organizational approaches emphasize discovery, visibility, risk classification, governance, and remediation, this model operates immediately before disclosure and can consume their outputs, including service inventories, approved alternatives, restrictions, and assurance information.

The evaluation does not show that deployment reduces Shadow AI; that behavioural effect requires empirical study. It shows only a mechanism for surfacing governance information at the decision point before disclosure.

The structural pattern may transfer beyond higher education, but cross-sector generalizability is not established and would require local policy, data mapping, service-assurance, and escalation rules.ΠD

6.4 Limitations and future validation

Five limitation families qualify the findings.

First, input quality matters. Incorrect data classification, omitted protection tags, or a misdescribed purpose can prevent relevant rules from being triggered. Second, governance quality matters. Incomplete, outdated, inconsistent, overly permissive, or overly restrictive policy can generate inappropriate recommendations even when the formal decision logic is correct. VERIFY can prevent policy silence from becoming implicit permission, but the model does not itself resolve conflicting institutional rules. Third, service information is time-dependent. Provider terms, retention, training or reuse arrangements, subprocessors, configurations, contractual protections, and institutional approvals may change. Fourth, decision scope is bounded. De-identification may be incomplete; decisions are valid only for the assessed configuration; and multiple individually acceptable prompts may cumulatively reveal a more sensitive context. Autonomous downstream actions by AI agents remain outside the present scope. Fifth, validation is analytical rather than

empirical. The study does not establish usability, user compliance, incident reduction, organizational adoption, legal correctness across jurisdictions, or cross-sector generalizability.

The next research stage should therefore test the model empirically. Controlled user studies could examine whether users identify D, S, and P correctly, understand the six actions, interpret institutional policy more consistently, and avoid treating ALLOW as an unconditional guarantee. Institutional pilots could test service registries, policy mappings, escalation pathways, and explanation traceability. Further work should also address cumulative disclosure across multi-turn interactions and authorization of agentic actions.

7. Conclusions

This study addresses the next action a higher-education user should take before submitting institutional information to a GenAI service. As formalized in Eq. (2), the proposed model combines data characteristics, service assurance and approved scope, intended use, contextual facts, and institutional policy to determine the next pre-submission action.

The research question is answered by structuring the point-of-use decision as a policy-grounded assessment of D, S, P, and C, yielding one of six next actions under Π. Rather than returning only allow/deny, the model distinguishes six actionable outcomes: proceed, transform and reassess, use an approved service, seek review, do not submit, or verify missing information. Across 20 constructed higher-education scenarios, the model reproduced the analytically derived reference action, responded coherently to controlled changes in data, service, and purpose, and preserved boundaries among uncertainty, incompatibility, review, and prohibition. These findings support internal analytical consistency under the reference policy profile, not independent empirical effectiveness.

The contribution is deliberately narrower than general Shadow AI or university AI governance. Existing governance remains responsible for policy, classification, service assurance, approved scope, and escalation; the model supplies a compact user-facing layer that translates a concrete configuration into the next policy-consistent action while preserving uncertainty and reassessment.

Future work should test comprehension, policy-interpretation consistency, service registries, cumulative disclosure, and agentic extensions through user studies and institutional pilots.

R E F E R E N C E S

[1] Y. An, J. H. Yu and S. James, Int. J. Educ. Technol. High. Educ. 22, 10 (2025).

[2] OECD, OECD Digital Education Outlook 2026 (OECD Publishing, Paris, 2026).

[3] OECD, OECD Education Spotlights 23, responsible GenAI adoption in higher education (2026).

[4] M. Li et al., Studies in Higher Education, advance online publication (2026).

[5] H. N. Nguyen, MethodsX 17, 104063 (2026).

[6] B. Y. Ng et al., Proc. Privacy Enhancing Technol. 2026(4), 1055-1071 (2026).

[7] M. Silic, D. Silic and K. Kind-Trüller, Strategic Change 34(5), 291-304 (2025).

[8] A. Zentner and T. West, Shadow AI in Higher Education, SSRN (2026).

[9] G. Sebastian, Technol. Forecast. Soc. Change 229, 124697 (2026).

[10] T. Cherner, T. S. Foulger and M. Donnelly, TechTrends 69(1), 84-99 (2025).

[11] S. Al Mandalawi et al., Proc. KSE 2025, IEEE (2025).

[12] M. Silic and A. Back, Computers & Security 45, 274-283 (2014).

[13] S. Klotz et al., Int. J. Inf. Syst. Project Manag. 7(1), 15-43 (2019).

[14] D. Fürstenau, H. Rothe and M. Sandner, Bus. Inf. Syst. Eng. 63, 97-111 (2021).

[15] S. G. Orr et al., Cryptologia 48(1), 26-90 (2024).

[16] A. Ojowa, M. Marques and A. Goncalves, Shadow AI in Organisations, Preprints.org (2026).

[17] NIST, Generative Artificial Intelligence Profile, NIST AI 600-1 (2024).

[18] ISO/IEC 42001:2023, Artificial intelligence - Management system (ISO/IEC, 2023).

[19] F. Miao and W. Holmes, Guidance for Generative AI in Education and Research (UNESCO, 2023).

[20] Stanford University, Responsible AI at Stanford, uit.stanford.edu (2026).

[21] University of Oxford, Policy for using generative AI in research, ox.ac.uk (2026).

[22] University of Oxford, Use Generative AI services safely, infosec.ox.ac.uk (2026).

[23] Wharton School, Generative AI resources, support.wharton.upenn.edu (2026).

[24] European Union, AI Act, Regulation (EU) 2024/1689, amended by Regulation (EU) 2026/1744.