How Are U.S. Universities Responding to AI? An Audit of Governance Capacity

ANONYMOUS AUTHOR(S)

Universities are rapidly adopting artificial intelligence in teaching and assessment, yet there is limited empirical visibility into how these institutions publicly govern its use. We present ACAI-US79, an institutional audit of AI governance across 79 U.S. universities, and ACAI, the Academic AI Capacity Index—an interpretable measure of publicly articulated governance capacity. The audit evaluates four domains—policy clarity, faculty support, feedback mechanisms, and AI detection tool governance—using time-bounded review of institutionally authoritative materials. ACAI does not assess technical capability or ethical intent; it measures the public legibility of institutional structures that allocate authority and accountability. We observe substantial variation in AI governance capacity across institutions, with recurring gaps in procedural safeguards and feedback mechanisms. Governance capacity does not consistently track research intensity: institutions with extensive AI research activity do not necessarily articulate stronger governance frameworks. We release the dataset, audit instrument, and public website at http://acai-us79.org/ to support transparency, critique, and institutional self-reflection, contributing to increased organizational accountability.

CCS Concepts: • **Applied computing** → **Education**; **Law**; *Computer-managed instruction.*

Additional Key Words and Phrases: AI, university, policy, governance

ACM Reference Format:
Anonymous Author(s). 2026. How Are U.S. Universities Responding to AI? An Audit of Governance Capacity. In *Proceedings of 2026 ACM Conference on Fairness, Accountability, and Transparency (FAccT '26).* ACM, New York, NY, USA, 33 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

Artificial intelligence (AI) is increasingly embedded into universities, shaping teaching, assessment, research, and administration [1, 2, 6, 10, 18, 21, 22, 42, 54, 55]. These deployments raise concerns about fairness, accountability, and harm, yet there is limited empirical evidence about how universities govern AI use in practice. Existing evaluations focus largely on technical systems [15], individual university responses [48], or national capacity (see Appendix F.2), *leaving a critical gap in understanding the institutional infrastructures through which AI use is authorized, constrained, contested, revised, and supported within higher education.* Moreover, U.S.-focused studies that systematically examine institutional AI policies have largely centered on top-ranked or R1 universities [31, 56, 58], obscuring variation across the broader higher-education landscape.

This gap matters for accountability. Universities exercise significant power over students and faculty,[1] and AI-related governance decisions – such as the use of AI detection tools in academic integrity enforcement [27, 51, 59] (more in Appendix G), faculty discretion over permissible AI use, or access to appeals – can have material consequences for equity and due process. Despite these stakes, there is no systematic, reproducible method for auditing how universities

[1]We use *faculty* as an umbrella term for all instructional staff, including non-tenure-track faculty. We note that this label encompasses roles with differing levels of authority, security, and participation in institutional governance, which vary across universities.

Manuscript submitted to ACM

publicly articulate AI governance commitments. As a result, claims about 'responsible AI' in higher education often lack an empirical foundation, risking both weak accountability and conceptual dilution of responsibility itself in the context of AI governance.

In this paper, we present **ACAI-US79**, an institutional audit dataset capturing publicly articulated AI governance practices across 79 diverse U.S. universities. From this dataset, we derive **ACAI**, the **Ac**ademic **A**I Capacity **I**ndex, an interpretable index that aggregates audit findings to produce a structured ranking of institutions based on the public legibility of formal policies, resources, and oversight mechanisms. This ranking is not a judgment of ethical adequacy, institutional intent, or internal practice; rather, it reflects differences in what governance artifacts are publicly visible under a consistent audit protocol. We release the dataset, annotation schema, audit toolkit, and public interface to support transparency, critique, and reuse, providing a baseline view of how universities currently articulate accountability around AI through publicly accessible governance materials.

We explicitly invite alternative reuse and contestation of ACAI. Institutions may reasonably dispute individual annotations, weighting choices, or domain boundaries, and such disagreement should be treated as a productive extension of the audit rather than a failure of the framework. Because ACAI relies exclusively on publicly available materials and a reproducible protocol, it is designed to support re-audits, counter-audits, and institutional self-assessment over time. In this sense, ACAI is not a static measurement, but an infrastructure for ongoing revision, critique, and accountability grounded in the public legibility of institutional governance.

To operationalize this audit, we focus on publicly legible institutional signals, and we assess AI governance capacity across four governance domains: A. Policy Clarity, B. Faculty Support, C. Feedback Loops, and D. Detection Tools. Annotators conduct time-bounded reviews of institutionally authoritative materials to evaluate whether relevant governance mechanisms are present, partially specified, absent, or conflicting. These assessments are aggregated into ACAI, yielding a transparent, diagnostic index that enables cross-institutional comparison without conflating governance capacity with technical expertise.

Lower ACAI scores should not be read as failures of responsibility or institutional care. Rather, they indicate that governance mechanisms are less publicly specified or harder to locate, which presents a distinct accountability risk regardless of intent. While publicly available materials do not capture internal deliberations or informal practices, they constitute the primary means through which universities communicate authority, procedural expectations, and avenues for recourse. For this reason, we treat public legibility as a necessary condition for accountability and as an appropriate object of empirical audit. Reliance on informal or "word-of-mouth" governance is not neutral: it systematically advantages actors with greater institutional access while disadvantaging students and faculty who must rely on publicly articulated rules. Publicly legible governance therefore establishes a minimally equitable baseline and is a prerequisite for procedurally just institutional AI governance.

We make the following contributions:

(1) An Open Dataset, Toolkit, and Website: We release ACAI-US79, a publicly available dataset for auditing AI governance across 79 U.S. universities, along with a reproducible annotation schema and audit toolkit (available at https://anonymous.4open.science/r/ACAI-3D27), and public website (available at https://acai-us79.org/).

(2) The Academic AI Capacity Index (ACAI): We introduce ACAI, an interpretable index for evaluating the public legibility of institutional AI governance, grounded in principles of accountability and procedural justice.

(3) An Empirical Institutional Audit: We apply ACAI in the first large-scale, reproducible audit of AI governance in U.S. higher education, demonstrating that accountability gaps persist even among leading AI research institutions.

Annotation Guidelines

You will review publicly available web pages for <UNIVERSITY> to determine whether specific AI-related policy statements are addressed by the institution.

Use ONLY the links provided below and any pages, sections, PDFs, or subpages that are directly reachable by clicking links on those pages (e.g., menus, internal links, or document links). Do NOT use external search engines or sources found from outside this list: <LINKS>.

Evaluate each statement independently. Spend no more than 5 minutes per statement.

For each statement:

1. Select exactly one classification:

- *Present/Yes* — A clear statement directly addressing the item is found on an institutional page within 5 minutes.
- *Partial/Implicit/Somewhat* — The item is mentioned or implied, but key details are missing.
- *Absent/No* — You reasonably searched the allowed sources and did not find relevant content.
- *Unclear or Took Longer Than 5 Minutes* — Navigation difficulty, vague language, or time limits prevented a confident decision.
- *Conflicting Information* — Different institutional sources provide contradictory guidance for the same item.

2. Provide the most relevant URL(s) from the allowed sources that support your selection. If you selected Absent or Unclear, provide the main page(s) you checked.

Statements Organized by Governance Domain:

A. POLICY CLARITY — *Policies defining institutional expectations, terminology, and academic integrity adaptations.*

A1. The university defines "AI use," "AI assistance," or "AI-generated content."
A2. The university defines standards for citing AI-generated material.

B. FACULTY SUPPORT — *Resources that enable faculty to integrate, regulate, or teach with AI.*

B1. The university provides guidance, training, or resources for faculty on AI-related teaching practices.
B2. Official examples of appropriate and/or prohibited AI use are provided (e.g. example AI use cases, example prompts).
B3. A faculty committee or group focused on teaching and learning about AI exists.
B4. Faculty are offered syllabus language examples (e.g. use AI/don't use AI/selectively use AI).

C. FEEDBACK LOOPS — *Mechanisms through which universities gather input, revise policies, and communicate decisions.*

C1. A faculty committee or advisory group focused on university AI policy or governance exists.
C2. A student committee or advisory group focused on university AI policy or governance exists.
C3. The university publishes AI policy update logs or explains revisions.

D. DETECTION TOOLS — *Institutional stance toward AI detection technologies.*

D1. The university restricts, discourages, or warns against the use of AI detection tools.
D2. Student misconduct determinations require human review and cannot be based solely on AI detection tools.

Fig. 1. **Annotation Instructions for ACAI Calculation, Organized By Governance Domain:** ACAI calculation details are provided in §2.2, governance domains are detailed in §2.2.1, and additional annotation details are in Appendix C.3.

(4) A Critical Evaluation of Automated Approaches to Governance Auditing: We evaluate whether large language models can approximate human governance judgments, showing that current automated approaches produce unstable and misleading institutional rankings.

(5) Actionable Recommendations for Universities: We translate our findings into concrete recommendations for strengthening institutional capacity for accountable AI governance, particularly with respect to feedback, review, and procedural safeguards.

By focusing on institutional governance rather than technical capability, this work advances organizational accountability. Institutional audits make visible the structures through which power is exercised, providing an empirical foundation for more transparent and procedurally just AI governance in higher education.

2 An Audit of Institutional Capacity

We first clarify what we mean by institutional capacity and why public legibility provides an appropriate object of empirical audit. This section conceptualizes AI governance in higher education as a question of institutional capacity and accountability, not technical sophistication or ethical aspiration. We treat universities as powerful organizational actors that structure how AI systems are authorized, constrained, and contested through policies, procedures, and oversight mechanisms. Adopting an audit perspective, we examine the public legibility of these governance arrangements: what is made visible, enforceable, and contestable to students, faculty, and other affected parties. This framing motivates our audit design and the construction of ACAI as a tool for evaluating how governance capacity is institutionalized across universities.

2.1 ACAI-US79: A Benchmarking Dataset

Our annotation protocol – shown in Figure 6 – operationalizes institutional capacity as publicly legible, time-bounded, and normatively specific governance, rather than as the mere presence of AI-related content. Capacity is understood as what institutions make visible and actionable to affected stakeholders within reasonable time and effort constraints. The empirical focus on 79 U.S. universities is motivated by the country's prominent position in global AI capacity indices [15, 32, 34, 38], detailed in Appendix F.2.

University Selection. The universities selected for the ACAI-US79 dataset and their corresponding attributes are shown in Table 1. The dataset was constructed using a purposive, diversity-oriented selection strategy designed to surface variation in how AI governance is institutionalized across U.S. higher education. The selection emphasizes institutional heterogeneity along dimensions that shape authority, accountability, and public legibility.

Within each U.S. Census region [53] — South, West, Midwest, and Northeast – we sought to include institutions spanning research intensity and institutional organization. Specifically, for each region we targeted approximately four institutions in each of the following categories: Public Research universities (R1 or R2), Private Research universities (R1 or R2), and Teaching/Liberal Arts colleges, categorized according to the Carnegie Classifications [4]. The choice of four institutions per category per region was a pragmatic design decision rather than a theoretical threshold, balancing between breadth, depth, and cost feasibility in the audit. Institution size was not used as an explicit stratification variable, thus the size distribution in ACAI-US79 emerges from the selection process and size is therefore treated analytically as a contextual attribute. We split size by tertiles into Small, Medium, and Large for comparison. Overall, this structure was intended to capture *differences in organizational incentives* that plausibly affect how AI governance is articulated.

Link Retrieval. For each university included in the audit, we systematically collected institutionally authoritative documents and official web links[2] across seven recall-oriented categories, with examples shown in Appendix C.2. These categories correspond to distinct organizational surfaces through which governance is commonly articulated [58], initialized by a manual author review of a small set of university policies, and then iterated on during our search process which is detailed in Appendix C.

Hence, the following categories serve as search lenses: (1) University Policies or Guidelines: University-level policies addressing technology use, data governance, research ethics, or academic administration; (2) Center for Teaching & Learning: Centrally maintained guidance for faculty on pedagogical use of AI, syllabus adaptation, and instructional support;

[2]Collection of links and annotation occurred from 12/8-12/30/25.

(3) AI Institute/Initiative/Center: Institutionally recognized units focused on AI, digital ethics, or related governance-relevant coordination; (4) Library Guide: Public-facing library resources addressing generative AI, citation practices, or responsible research use; (5) Academic Integrity: Policies or guidance governing the relationship between AI, authorship, plagiarism, and assessment; (6) AI Committee: Formal committees or task forces charged with evaluating or coordinating institutional responses to AI; and (7) Other Relevant Links: Additional institutionally authoritative materials relevant to AI governance.

The audit does not assume that governance capacity is limited to legally binding rules. Instead, we operationalize capacity as institutionalization: the presence of standing reference points – such as policies, offices, committees, or officially maintained resources – that persist over time and orient behavior by establishing expectations about authority, coordination, and acceptable practice. They represent *ongoing points of reference that a student, instructor, administrator, member of the press, or member of the public could reasonably consult* to understand how AI is governed at the institution. In contrast, transient communications such as news articles, announcements, or blog posts were excluded.[3] While such materials may signal institutional intent or activity, they do not establish durable roles, procedures, or accountability structures. Including them would collapse institutionalization into communication and systematically overstate governance capacity. Importantly, the links collected through this process are treated as candidate surfaces of governance, not as governance determinations themselves. As shown in Figure 1, annotators were instructed to use these materials as evidence when evaluating whether specific governance statements were present, partial, or absent. We provide a detailed description of our link retrieval procedure in Appendix C.

Human Annotation via Prolific. All annotation was conducted by paid human annotators recruited via the Prolific[4] platform, following the annotation instructions shown in Figure 1. Annotators labeled only publicly available institutional materials under standard compensated microtask conditions, and no PII or sensitive data were collected.[5] Tasks were estimated to take approximately 30 minutes and were compensated at \$6.00, consistent with Prolific's recommended compensation. Each university was independently annotated by three annotators, administered via an external survey instrument. Participation was restricted to U.S.-based, English-fluent annotators using desktop or laptop devices, enforced through Prolific's prescreening tools. To preserve the audit's focus on public legibility, annotators were instructed not to use external search engines or prior knowledge, and each item was evaluated under a strict time limit. To support data quality, we enabled Prolific's "reject exceptionally fast submissions" safeguard, which automatically flags submissions completed at implausibly short durations relative to the estimated task time, helping to filter inattentive or agent-driven responses. To assess alignment between expert and Prolific annotations, two computational PhD students independently annotated a subset of 10 universities using the same audit protocol. Aggregate agreement between the expert and Prolific annotations was moderate to strong (Pearson $r = 0.56$, Spearman $\rho = 0.57$, Kendall $\tau = 0.49$; all $p < 0.05$), with Prolific scores exhibiting a small positive bias towards higher ratings (+0.10 on a 0–1 scale). Further, we compare our human annotated results to an LLM-driven audit, detailed in §2.4.

2.2 ACAI: The Academic AI Capacity Index

To enable systematic comparison of governance capacity across institutions, we construct **ACAI**, the **Ac**ademic **A**I Capacity **I**ndex. While the index necessarily collapses nuances present in more detailed datasets, ACAI serves as a first

[3]Prior studies [31, 56, 58] can be referenced for details on these communication types.
[4]Prolific is a crowdworker platform commonly used for academic research and data annotation tasks: https://www.prolific.com/data-annotation
[5]We provide the privacy policy in Figure 8.

step toward understanding the AI governance landscape of U.S. universities. ACAI aggregates governance indicators across four domain (Figure 1) into a single interpretable score:

$$\mathrm{ACAI}_u = \frac{\sum_{d\in\{A,B,C,D\}} w_d \left(\frac{100}{J_d}\sum_{j=1}^{J_d}\frac{1}{k}\sum_{i=1}^{k} I_{d,u,i,j}\right)}{\sum_{d\in\{A,B,C,D\}} w_d} \tag{1}$$

where $I_{d,u,i,j} \in \{0, 0.5, 1\}$ is the score assigned by annotator i to indicator item j within governance domain d – where $d \in$\{A,B,C,D\}: A. Policy Clarity, B. Faculty Support, C. Feedback Loops, D. Detection Tools – at university u, and J_d is the number of indicators in governance domain d. Indicator scores are coded as 1.0 (*Present/Yes*), 0.5 (*Partial/Implicit/Somewhat*), and 0.0 (*Absent/No*, *Unclear or Took Longer Than 5 Minutes*, or *Conflicting Information*). Domain scores are computed as the mean across indicators and annotators, and then aggregated via a weighted sum and scaled to $[0, 100]$. Higher values indicate greater publicly articulated AI governance capacity. Weighting coefficients, w_d, encode normative priorities about which governance functions matter most. To ensure findings are not artifacts of these choices, we assess robustness under alternative weighting schemes, varying $1 \leq w_d \leq 4$ (see §section F5 and §Appendix D for details).

A note on aggregation: We intentionally aggregate indicator scores by averaging across annotators rather than resolving disagreement through expert adjudication or majority vote. This audit principle reflects the audit's focus on public legibility rather than institutional intent, internal consistency, or expert interpretation. ACAI is explicitly not designed to capture how governance materials might be interpreted by legal counsel, administrators, or domain experts, but how they are encountered by external readers operating under realistic time and access constraints. When multiple annotators reviewing the same publicly available materials arrive at different judgments about whether a governance mechanism is present, partial, or absent, that variation is treated as an empirical signal of ambiguity in the underlying institutional artifacts. Hence, annotator disagreement is evidence about how clearly governance is articulated [3]. Averaging across multiple annotations aligns with the audit framing of ACAI as a measure of publicly articulated governance capacity: governance that requires expert interpretation or insider knowledge to interpret functions as weaker governance in practice, regardless of internal deliberation or intent. ACAI thus measures how governance is *encountered* by external readers under realistic constraints, not how it might be interpreted by insiders. Because the index relies on publicly available materials, it should be interpreted as a lower bound on institutional capacity.

2.2.1 Governance Domains. We now turn to the four governance domains covered in our study:

A. Policy Clarity. This domain captures the extent to which universities publicly articulate clear, institution-level expectations regarding AI use in academic contexts, following a long legal tradition establishing the importance of clear definitions to support stable legal interpretation [19, 28]. Policy clarity includes the definition of key terms (e.g., "AI use," "AI assistance," or "AI-generated content"), guidance on attribution or citation of AI-generated material, and the adaptation of existing academic integrity frameworks to account for AI-mediated authorship. Prior work has shown heterogeneity in how universities define and communicate AI-related expectations, with many institutions relying on vague or decentralized guidance [31, 56]. In the absence of clear, publicly legible policy language, responsibility for interpreting acceptable AI use is often shifted to individual faculty or students, increasing the risk of inconsistent enforcement and inequitable outcomes [41]. Policy clarity therefore functions as a foundational component of institutional governance capacity, establishing shared reference points for authority, compliance, and contestation.

ACAI Rank	CSRankings_{AI} Rank	Institution	Type	Research Activity	Region	Size	ACAI Score
1	105	University of New Hampshire	Public Research	R1	Northeast	Medium	81.82
2	96	Portland State University	Public Research	R2	West	Large	80.30
3	8	Stanford University	Private Research	R1	West	Medium	80.30
4	14	University of Texas at Austin	Public Research	R1	South	Large	77.27
5	42	University of Notre Dame	Private Research	R1	Midwest	Medium	75.76
6	137	Baylor University	Private Research	R1	South	Large	74.24
7	44	University at Buffalo	Public Research	R1	Northeast	Large	74.24
8	79	University of Florida	Public Research	R1	South	Large	71.21
9	9	University of Michigan at Ann Arbor	Public Research	R1	Midwest	Large	71.21
10	*un.*	Rowan University	Public Research	R2	Northeast	Large	71.21
11	29	Stony Brook University	Public Research	R1	Northeast	Large	71.21
12	*un.*	Lewis & Clark College	Teaching/Liberal Arts	–	West	Small	69.70
13	7	University of California, Berkeley	Public Research	R1	West	Large	69.70
14	27	Texas A&M University	Public Research	R1	South	Large	69.70
15	91	Case Western Reserve University	Private Research	R1	Midwest	Medium	69.70
16	*un.*	Lafayette College	Teaching/Liberal Arts	–	Northeast	Small	69.70
17	*un.*	California State University, Long Beach	Public Research	R2	West	Large	68.18
18	25	University of North Carolina at Chapel Hill	Public Research	R1	South	Large	68.18
19	6	Cornell University	Private Research	R1	Northeast	Medium	68.18
20	117	Brandeis University	Private Research	R1	Northeast	Small	68.18
21	169	Southern Methodist University	Private Research	R1	South	Medium	68.18
22	*un.*	Chapman University	Private Research	R2	West	Medium	68.18
23	*un.*	Howard University	Private Research	R1	South	Medium	68.18
24	81	University of South Florida	Public Research	R1	South	Large	66.67
25	100	Syracuse University	Private Research	R1	Northeast	Large	66.67
26	*un.*	University of Wyoming	Public Research	R1	West	Medium	65.15
27	37	The Ohio State University	Public Research	R1	Midwest	Large	65.15
28	15	University of Southern California	Private Research	R1	West	Large	65.15
29	*un.*	Mercer University 9	Private Research	R2	South	Medium	63.64
30	142	DePaul University	Private Research	R2	Midwest	Large	63.64
31	29	Arizona State University	Public Research	R1	West	Large	63.64
32	*un.*	Northern Illinois University	Public Research	R2	Midwest	Medium	63.64
33	*un.*	University of South Alabama	Public Research	R2	South	Medium	63.64
34	*un.*	Fordham University	Private Research	R2	Northeast	Medium	63.64
35	169	Florida Institute of Technology	Private Research	R2	South	Small	63.64
36	*un.*	Illinois State University	Public Research	R2	Midwest	Large	62.12
37	*un.*	Pepperdine University	Private Research	R2	West	Medium	60.61
38	46	University of Chicago	Private Research	R1	Midwest	Large	60.61
39	*un.*	Montclair State University	Public Research	R2	Northeast	Medium	60.61
40	100	Binghamton University	Public Research	R1	Northeast	Medium	60.61
41	*un.*	Lake Forest College	Teaching/Liberal Arts	–	Midwest	Small	60.61
42	79	Iowa State University	Public Research	R1	Midwest	Large	57.58
43	*un.*	Carleton College	Teaching/Liberal Arts	–	Midwest	Small	57.58
44	16	University of Washington-Seattle	Public Research	R1	West	Large	57.58
45	*un.*	San José State University	Public Research	R2	West	Large	57.58
46	*un.*	Southern University and A & M College	Public Research	R2	South	Small	56.06
47	*un.*	Colby College	Teaching/Liberal Arts	–	Northeast	Small	56.06
48	*un.*	Skidmore College	Teaching/Liberal Arts	–	Northeast	Small	56.06
49	*un.*	Saint Louis University	Private Research	R1	Midwest	Medium	56.06
50	*un.*	Reed College	Teaching/Liberal Arts	–	West	Small	56.06
51	*un.*	Southern Wesleyan University	Teaching/Liberal Arts	–	South	Small	54.55
52	*un.*	Davidson College	Teaching/Liberal Arts	–	South	Small	54.55
53	*un.*	University of Colorado Colorado Springs	Public Research	R2	West	Medium	54.55
54	2	University of Illinois Urbana-Champaign	Public Research	R1	Midwest	Large	54.55
55	142	Wichita State University	Public Research	R2	Midwest	Medium	53.03
56	*un.*	Wofford College	Teaching/Liberal Arts	–	South	Small	51.52
57	*un.*	Georgia Southern University	Public Research	R2	South	Large	51.52
58	*un.*	University of Denver	Private Research	R1	West	Medium	51.52
59	*un.*	Westminster University	Teaching/Liberal Arts	–	West	Small	50.00
60	*un.*	Ball State University	Public Research	R2	Midwest	Large	48.48
61	*un.*	Rhodes College	Teaching/Liberal Arts	–	South	Small	48.48
62	49	Brown University	Private Research	R1	Northeast	Medium	48.48
63	*un.*	Clark University	Private Research	R2	Northeast	Small	48.48
64	*un.*	Abilene Christian University	Private Research	R2	South	Small	48.48
65	*un.*	Wesleyan University	Teaching/Liberal Arts	–	Northeast	Small	46.97
66	100	Illinois Institute of Technology	Private Research	R2	Midwest	Small	46.97
67	142	Nova Southeastern University	Private Research	R1	South	Large	46.97
68	*un.*	Kean University	Public Research	R2	Northeast	Medium	45.45
69	*un.*	Occidental College	Teaching/Liberal Arts	–	West	Small	43.94
70	61	Stevens Institute of Technology	Private Research	R2	Northeast	Small	43.94
71	71	California Institute of Technology	Private Research	R1	West	Small	43.94
72	*un.*	Long Island University	Private Research	R2	Northeast	Medium	42.42
73	*un.*	Marquette University	Private Research	R2	Midwest	Medium	40.91
74	*un.*	Beloit College	Teaching/Liberal Arts	–	Midwest	Small	39.39
75	*un.*	Grinnell College	Teaching/Liberal Arts	–	Midwest	Small	34.85
76	*un.*	Jackson State University	Public Research	R2	South	Small	31.82
77	*un.*	Clark Atlanta University	Private Research	R2	South	Small	28.79
78	*un.*	Creighton University	Private Research	R2	Midwest	Medium	27.27
79	*un.*	University of Massachusetts at Dartmouth	Public Research	R2	Northeast	Small	27.27

Table 1. **Institutional Rankings for ACAI-US79 with ACAI and CSRankings_{AI}**: Details for ACAI calculations in §2; details for *CSRankings_{AI}* in Appendix E; *Region* is classified based on the U.S. Census [53]; *Type* and *Research Activity* are classified based on the Carnegie Classifications [4]; *Size* is split based on tertile buckets.

B. Faculty Support. This domain reflects the extent to which universities provide institutionally maintained resources that enable faculty to engage with AI in teaching and assessment in informed and supported ways – ultimately to support student learning [30]. Faculty support includes guidance on pedagogical uses of AI, training opportunities, example use cases or prohibitions, model syllabus language, AI professional development opportunities, and the presence of faculty-focused committees or working groups concerned with AI and teaching [31, 41, 56]. Faculty are frequently expected to make consequential decisions about AI use – such as whether and how to permit AI in coursework – without adequate institutional support or coordination [31, 41]. The presence of formal faculty support mechanisms signals that responsibility for AI governance is not delegated entirely to individual faculty, but is instead recognized also as an institutional obligation requiring shared infrastructure and expertise. Importantly, faculty support directly impacts students: clear guidance helps faculty communicate consistent expectations to students, while faculty committees and resources provide the foundation for student-facing policies articulated in the other domains.

C. Feedback Loops. This domain reflects concerns articulated in prior survey research and aligns with recommendations from the July 2025 American Association of University Professors (AAUP) report *Artificial Intelligence and Academic Professions* [41]. The report draws on approximately 500 survey responses from AAUP members regarding their experiences with AI and other educational technologies. A central finding was a widespread concern (reported by 71% of respondents) about the disconnect between administrative decision-making on AI policy and meaningful faculty and student input. In response, the report recommends the adoption of "meaningful shared governance policies and practices," including committees composed of faculty, staff, and/or students, as well as increased transparency around AI-related decisions and policy changes. More broadly, scholarship on algorithmic governance emphasizes that accountability requires institutionalized mechanisms for participation, feedback, and revision over time. Absent such mechanisms, governance frameworks risk functioning as static or symbolic commitments rather than durable, contestable structures of authority [33, 46]. For an extended discussion on this subdomain, see Appendix F.1.

D. Detection Tools. This domain captures institutional stances toward AI detection technologies used in academic integrity enforcement [51, 59]. Indicators in this domain assess whether universities restrict, discourage, or explicitly govern the use of AI detection tools, and whether procedural safeguards – such as requirements for human review – are articulated [26]. Prior research has documented significant technical limitations and bias in AI detection systems, as well as their potential to produce false positives with serious consequences for students [27, 56]. Despite these risks, institutional guidance on detection tools is often limited, ambiguous, or silent on procedural constraints [31], potentially algorithmically shortcutting due process[6] protections [25, 43], and with little-to-no contestability [29]. Publicly articulated governance in this domain is therefore critical for clarifying the role of detection tools in decision-making, delineating authority between automated systems and human judgment, and protecting due process in academic misconduct determinations. For an extended discussion on this subdomain, see Appendix G.

The governance domains used in ACAI are not intended as an exhaustive or universal taxonomy. They reflect governance mechanisms that are currently most visible and auditable within U.S. higher education using publicly available materials and realistic time constraints. Governance domains may vary across national, legal, and institutional contexts, and future audits may adapt or expand this structure accordingly. In particular, as universities increasingly deploy AI monitoring [44] and auditing [42] systems (see §3.2), future iterations of ACAI may incorporate additional

[6]As noted by Khattak [25] on the U.S. context: *"As artificial intelligence systems increasingly assist decisionmaking in judicial and administrative processes, courts and administrative agencies face mounting pressure to merge innovation with legal tradition. These technologies are often praised for their efficiency. However, when the mechanisms by which they operate are impenetrable, they threaten to infringe upon core due process protections. The Constitution guarantees that individuals be informed of decisions affecting their rights, and to have a fair opportunity to contest those decisions in a meaningful way. When unclear algorithms replace human judgment, those guarantees are at risk of becoming procedural only by name."*

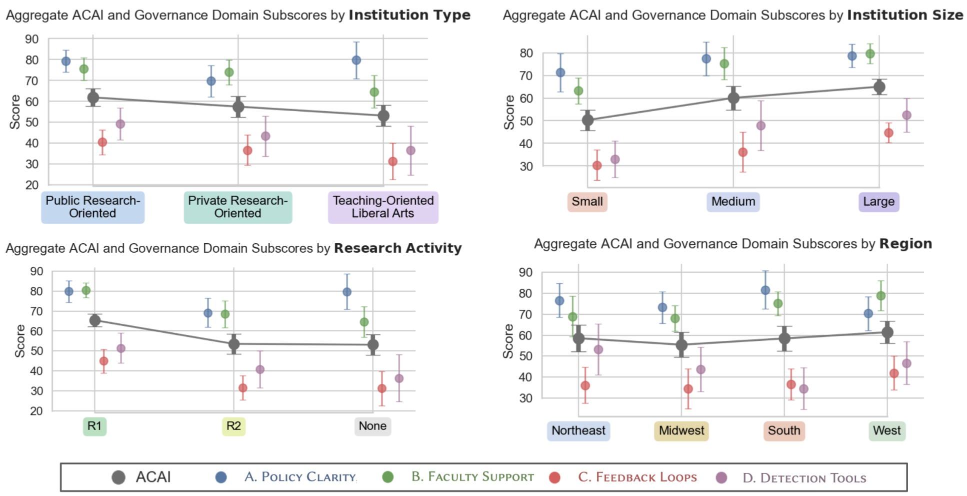


Fig. 2. **Aggregate ACAI and governance domain subscores (A,B,C,D) show that AI governance capacity varies with organizational structure, and that governance participation (C) and due process (D) are undersupported (▷F1, F2).**

governance domains to capture these practices. We therefore view ACAI as a flexible audit framework rather than a fixed index, designed to evolve alongside institutional AI governance.

2.3 Approximating AI Research Activity with CSRankings$_{AI}$

To contextualize institutional AI governance capacity relative to AI research activity, we used the rankings provided by CSRankings, a widely used, publicly available ranking of computer science research output shown in Figure 9. CSRankings aggregates publication counts across major computer science venues and allows filtering by research area and time period; for details on the specific configuration used in this study, see Appendix E. As shown in Table 1, we report the CSRankings$_{AI}$ rank of each ACAI-US79 university to facilitate our comparison of AI governance capacity and AI research activity, and examine correlations between these two rankings as shown in Figure 3. Notably, CSRankings only ranks research-active institutions by publication output; Teaching/Liberal Arts colleges and institutions without substantial computer science research activity do not appear in CSRankings and are therefore excluded from the correlation analysis in Figure 3.

2.4 LLM Study

Because universities increasingly rely on AI systems to assess, classify, and enforce academic norms [11, 16, 22, 44], we conducted an LLM study to examine whether similar systems can meaningfully evaluate institutional AI governance: we conducted an LLM study aligned with the ACAI-US79 audit framework. The prompt used is shown in Figure 11; we perform three independent runs for each sampling temperature $\tau \in \{0.5, 1.0, 1.5\}$. For each university-statement pair, the model was required to return a single categorical score and supporting URLs in a strictly validated JSON schema.

3 Findings

We use exploratory subgroup comparisons to examine how AI governance capacity is differentially institutionalized across higher-education contexts. Contrasts in ACAI scores across institutional type, research activity, region, and size

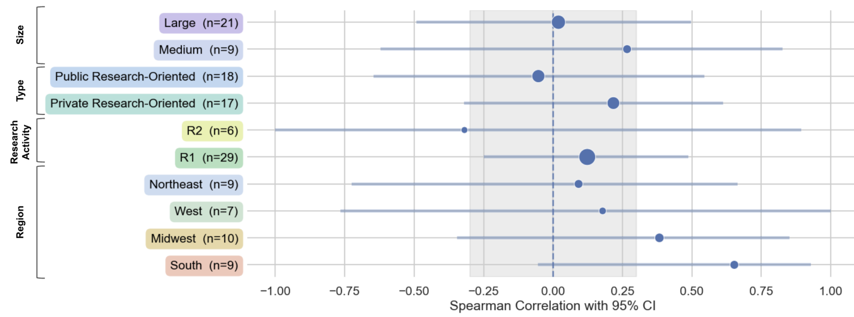


Fig. 3. **Spearman ρ rank correlations between ACAI ranks and CSRankings$_{AI}$ ranks across institutional subgroups show that AI governance capacity and AI research output are almost entirely uncorrelated (▷F3):** All CIs cross zero, unranked CSRankings$_{AI}$ universities are excluded, and categories with $n \leq 5$ are dropped. We provide detailed results in Table 7.

are used as a descriptive lens on organizational variation, with effect sizes and confidence intervals reported to convey the magnitude and uncertainty of observed differences. These subgroup comparisons are not treated as confirmatory evidence: the audit sample is purposive rather than probabilistic, subgroup categories are administratively defined and analytically coarse, and multiple overlapping contrasts are examined without correction. Accordingly, the results support interpretive claims about institutional accountability and structural incentives, rather than population-level inference, causal explanation, or claims of statistically significant subgroup differences.

▷ ***F1: AI governance capacity varies with organizational structure.*** As shown in Figure 2, Public Research universities tend to occupy higher positions in the ACAI distribution than Private Research or Teaching/Liberal Arts universities; larger institutions (Large and Medium) tend to exhibit higher ACAI scores than Small ones; R1 schools tend to show higher ACAI scores than R2 and unclassified (–) schools; and schools in the Western region tend to show slightly higher ACAI scores. These differences align with known variation in institutional oversight arrangements and coordination demands: public universities commonly operate under statutory or regulatory accountability frameworks and maintain centralized administrative infrastructures, which are visible in the form of institution-level policies, guidance pages, and standing committees. In contrast, institutions characterized by more decentralized organizational structures often rely on localized practices, which may be less consistently reflected in publicly accessible governance artifacts. Additionally, large institutions typically coordinate governance across a greater number of academic units, faculty, and students, which is reflected in the presence of centrally maintained and publicly legible reference points. Overall, even among the highest-capacity institutions, ACAI scores reveal substantial room for improavement. For example, UC-Berkeley – the highest-scoring institution among Large, Public Research, R1 universities in the West – achieves an ACAI score of 69.70, indicating approximately 30% of audited governace indicators were absent, partial, or unclear across the four domains.

Importantly, ACAI does not assess research quality, ethical commitments, or internal decision-making processes. It captures whether AI-related governance mechanisms are publicly articulated and institutionally maintained. From this perspective, lower ACAI scores should not be interpreted as evidence of weaker concern or expertise, but as indicative of different approaches to organizing and communicating governance. These findings highlight a structural tension: institutional arrangements that emphasize decentralization or flexibility may be less visible in public-facing governance artifacts, even when substantive internal practices are present. However, this lack of public legibility introduces a

τ	Pearson r	Spearman ρ	Kendall τ	Spearman 95% CI	Mean Δ	$\Delta_{\text{Bottom 25\%}}$	$\Delta_{\text{Top 25\%}}$
0.5	0.56^{***}	0.60^{***}	0.43^{***}	[0.42, 0.74]	15.39	[0, 5]	[21, 69]
1.0	0.52^{***}	0.54^{***}	0.38^{***}	[0.36, 0.71]	16.23	[0, 5]	[22, 70]
1.5	0.53^{***}	0.52^{***}	0.37^{***}	[0.30, 0.68]	17.34	[0, 7]	[24, 67]

Table 2. **Correlations between human-labeled ACAI ranks and LLM-labeled ACAI ranks indicate that LLMs only weakly approximate human interpretive judgment (▷F4)**: For institution i, the absolute rank gap is defined as $\Delta(i) = |rank_{Human}(i) - rank_{LLM}(i)|$. $\Delta_{\text{Bottom 25\%}}$ and $\Delta_{\text{Top 25\%}}$ report the minimum and maximum values of Δ_i among institutions in the lower and upper quartiles of the rank gap distribution. Spearman correlations additionally report bootstrap 95% confidence intervals.

specific accountability risk: reliance on informal guidance or word-of-mouth governance differentially advantages actors with greater institutional access, while disadvantaging students and faculty who must rely on publicly accessible rules and procedures. Publicly articulated governance capacity thus functions as a necessary condition for procedural accountability, establishing a minimally equitable baseline that does not depend on social transmission or insider knowledge.

▷ ***F2: AI governance capacity is concentrated in policy articulation rather than participation or process.*** Figure 2 reveals that across institutional types, sizes, research intensities, and regions, scores in A. Policy Clarity and B. Faculty Support are systematically higher than those in C. Feedback Loops and D. Detection Tools. This consistent structural pattern indicates that publicly articulated AI governance capacity is concentrated in domains oriented toward rule articulation and instructional guidance, rather than in mechanisms that enable participation, feedback, or procedural constraint.

Notably, this gap persists even among institutions with otherwise high aggregate ACAI scores. Large, Public Research, and R1 universities – while exhibiting higher overall governance capacity – still show pronounced deficits in feedback and detection tool governance relative to policy articulation. This pattern suggests that differences are primarily quantitative rather than qualitative: AI governance capacity scales with organizational resources, but its internal composition remains skewed toward static guidance rather than durable procedural safeguards. In this sense, higher-capacity institutions often extend the same governance model rather than adopting qualitatively different forms of participatory or process-oriented governance.

As shown in Figure 2, although AI detection tools are frequently referenced in academic integrity materials, explicit procedural guidance governing their use is rare. This aligns with the findings of Wang et al. [56], who found in their analysis that while 57% of universities in their dataset mentioned common tools, none explicitly recommended their use. Institutions often fail to specify whether detection tools are advisory or determinative, how results should be interpreted, what safeguards exist against error, or what recourse is available to affected students. This gap creates ambiguity around authority and enforcement, and risks inconsistent or discretionary application in practice, as expanded upon in Appendix G, directly harming students, and creating the potential for an adversarial relationship between students and faculty which is counterproductive to a healthy learning environment.

▷ ***F3: AI governance capacity is largely uncorrelated with AI research output.*** As shown in Figure 3, AI governance capacity (ACAI) does not strongly correlate with AI-specific research output (CSRankings_{AI}), falling in the gray range; this dispersion indicates that AI technical leadership alone does not reliably translate into strong, visible governance practices. This is further demonstrated in Table 1: The highest ACAI score is achieved by the University of New Hampshire, followed by Portland State University and Stanford University. Two of the top three institutions are Public Research universities, including one classified as R2, and three within the top 20 are Small institutions, two of which are solely Teaching/Liberal Arts colleges, illustrating that high levels of publicly articulated governance capacity are observed across a range of institutional types and research classifications. Across the full sample, several Public Research / R2 universities – such as Rowan University and California State University, Long Beach – appear in higher positions in the ACAI distribution than many Private Research / R1 universities with substantial AI research activity. Conversely, multiple Private Research / R1 universities – such as Brown University and the California Institute of Technology – appear in the lower half of the ACAI distribution despite significant contributions to AI scholarship. These contrasts indicate that AI research intensity and AI governance capacity are orthogonal dimensions of institutional capability.

▷ ***F4: LLMs only partially reproduce human governance judgments.*** As shown in Table 2, across temperatures, LLM-generated ACAI rankings exhibit moderate ordinal agreement with human judgments ($\rho = 0.52 - 0.60$), yet individual institutions are frequently misranked, with mean absolute rank errors of approximately 15-17 positions. Importantly, in the aggregate these errors are not symmetric. As shown by the quartile breakdowns, institutions in the upper quartile of the rank-gap distribution experience extreme misrankings of up to 67–70 positions – nearly inverting the relative ordering of affected institutions. Even the lower quartile exhibits nontrivial discrepancies (0-5 positions). This pattern indicates that LLM outputs are structurally unstable. Taken together, these results suggest that while LLMs may approximate coarse aggregate patterns, they fail to reliably reproduce the fine-grained, interpretive distinctions required for institutional governance audits. Governance evaluation depends on contextual reading, procedural inference, and judgment under ambiguity – capacities that are not robustly captured by current LLM-based approaches. As a result, automated audits risk introducing arbitrary or misleading institutional comparisons, showing the continued necessity of human-centered audit methodologies for evaluating publicly articulated governance capacity.

▷ ***F5: ACAI rankings are robust to weighting choices and individual annotators.*** To assess whether ACAI rankings are artifacts of either normative aggregation choices or variation in annotator judgments, we conducted a series of robustness analyses varying both domain weights and annotator inclusion. We first evaluated four weighting schemes: an indicator-weighted baseline (w_A=2, w_B=4, w_C=3, w_D=2), equal (w_A=1, w_B=1, w_C=1, w_D=1), policy-heavy (w_A=1, w_B=1, w_C=2, w_D=2), and teaching-heavy (w_A=1, w_B=2, w_C=1, w_D=1), using percentile ranks. Rankings were highly stable across weighting schemes (Spearman ρ = 0.93–0.99, Pearson r = 0.93–0.99). Mean maximum rank shifts were modest 16.28 ranks, with an interquartile range of 14.5-22.5 ranks and a maximum of 46 ranks; details in Appendix D.

We compute inter-annotator agreement using multiple complementary metrics. At the level of individual items, agreement is modest and variable (Krippendorff's $\alpha = 0.26$), reflecting the interpretive and normative nature of governance assessment and the heterogeneity of institutional documentation. Average pairwise agreement shows a similar pattern (mean = 0.48), indicating systematic but incomplete convergence among annotators. When indicators are aggregated at the governance domain-level, agreement improves (α = 0.30), suggesting that higher-level governance constructs are more consistently interpretable than individual policy statements. We also evaluated robustness to individual annotators using leave-one-annotator-out recomputation of ACAI scores under the indicator-weighted scheme. Across all three exclusions, recomputed scores remained strongly correlated with the full-annotator index (r = 0.83–0.87, ρ = 0.75–0.87), indicating that no single annotator systematically altered the relative ordering of high- or low-capacity institutions.

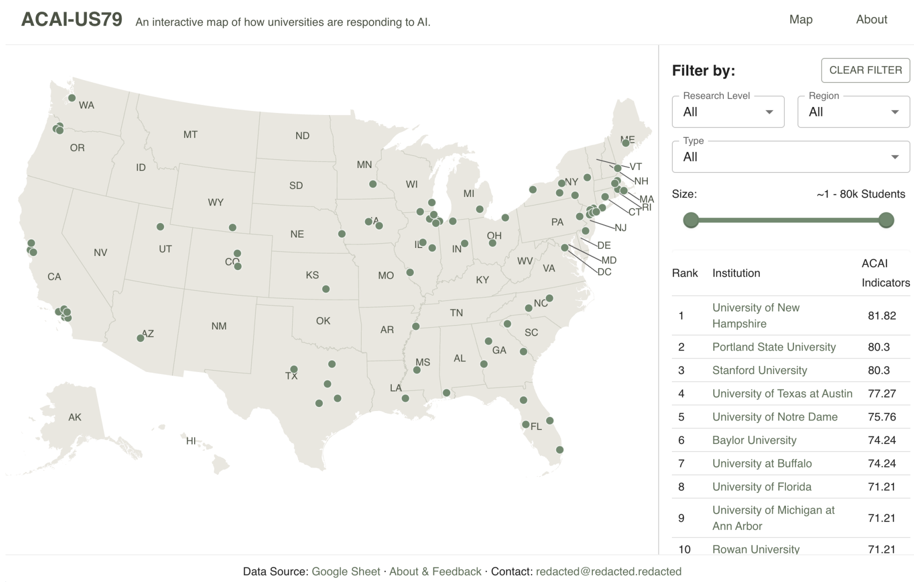


Fig. 4. **Interactive map of ACAI-US79 at https://acai-us79.org/**, visualizing the 79 U.S. universities and describing their publicly articulated governance capacity. Institutions are shown as clickable markers and ranked by ACAI score, with filters enabling comparison across research activity, institutional type, region, and size. Selecting an institution reveals its score and links to the publicly available policies, guidance, and governance materials reviewed in the audit, supporting traceability and independent inspection of how governance capacity is publicly articulated.

3.1 University Policy Recommendations

Based on the audit findings, we outline a set of policy recommendations aimed at strengthening institutional AI governance capacity in higher education.

▷ ***R1: Establish a centralized institution-level AI governance reference point (following ▷F1, F3).*** Universities should maintain clearly identifiable, centrally managed points of reference – such as policies, standing guidance pages, or designated offices – that articulate how AI use is governed. Reliance on informal norms or dispersed documentation makes governance difficult to locate and unevenly accessible to students and faculty.

▷ ***R2: Provide procedural clarity around academic integrity and AI detection tools (following ▷F1, F2).*** Where AI detection tools or integrity enforcement mechanisms are referenced, institutions should articulate clear procedures governing their use, limits, appeal processes, and responsible parties. Absent such guidance, detection practices risk being experienced as opaque, discretionary, or punitive.

▷ ***R3: Formalize feedback and revision mechanisms (following ▷F2).*** Governance capacity is strengthened when institutions specify how AI-related policies are reviewed, updated, and contested over time. Standing committees, task forces, or revision timelines signal that governance is ongoing rather than static or symbolic.

▷ ***R4: Treat public legibility as a core governance requirement (following ▷F1, F2, F3).*** Institutions should evaluate AI governance materials from the perspective of reasonable users – students, faculty, and administrators – who must locate and interpret guidance under time constraints. Governance that exists but is difficult to find or interpret functions as limited governance in practice.

3.2 Future Directions

Future work should examine how emerging forms of AI-mediated surveillance and labor automation jointly erode institutional governance capacity in higher education. Universities are rapidly deploying AI-based monitoring systems – such as AI detection tools (see §2.2.1 and Appendix G), fully automated proctoring based on behavioral analytics such as eye-tracking and click-tracking [44], fully automated verbal exams with voice AI [22] and LLM-as-a-judge grading [11, 16] – under the language of integrity and efficiency, yet these systems operate within asymmetrical power relations that render consent effectively coercive [12, 14, 45]. Students cannot *meaningfully opt out* without material penalty, while governance mechanisms lag behind technological adoption. At the same time, universities are reducing human interpretive labor through adjunctification while increasingly automating AI-mediated instruction and auditing [13, 42], displacing the very actors – teaching assistants, faculty, and staff – who translate policy into practice and provide critical feedback on institutional decisions. Together, these trends expand computational oversight while hollowing out human oversight, producing a net transfer of power from institutional governance to technical systems. Future research should develop participatory audit frameworks for educational AI that include students, faculty, and staff; foreground transparency, contestability, and review; and treat interpretive labor as a core governance infrastructure.

4 Conclusion

We present a large-scale institutional audit of publicly articulated AI governance in U.S. higher education, shifting attention from technical systems to the organizational infrastructures through which AI-related authority is exercised. Through ACAI-US79, a publicly released dataset of governance annotations across 79 U.S. universities, and the Academic AI Capacity Index (ACAI), we produce a comparative ranking of institutions based on the public legibility of their AI governance capacity. This ranking reveals substantial unevenness: governance capacity is frequently concentrated in rule articulation rather than in mechanisms for participation, feedback, or procedural safeguards – particularly around AI detection tools. Importantly, higher ACAI rankings do not consistently align with AI research intensity, indicating that institutional accountability is shaped more by organizational design and incentives than by technical leadership alone. While ACAI is a ranking, it is intended as a diagnostic rather than a normative judgment of institutional quality or ethical commitment: it captures what universities publicly formalize and communicate at a specific point in time, not internal deliberations, intent, or expertise. To support transparency, contestation, and longitudinal analysis, we publicly release ACAI-US79, the audit instrument, and an accompanying website that makes underlying governance artifacts directly inspectable and enables future re-audits as institutions update their policies. By rendering AI governance structures visible and comparable, this work provides an empirical foundation for studying institutional accountability and for advancing accountable AI governance in higher education.

5 Generative AI Usage Statement

We responsibly used AI technologies (ChatGPT, v5) in this paper to assist with search, and the styling and language of the writing, as well as code assistance.

References

[1] 2025. *EAAI-25: The 15th Symposium on Educational Advances in Artificial Intelligence.* Association for the Advancement of Artificial Intelligence (AAAI), Philadelphia, PA, USA. https://eaai-conf.github.io/year/eaai-25.html

[2] 2025. Teaching the Arts in the Age of A.I.: A Symposium. https://arts.umd.edu/teaching-arts-in-the-age-of-ai-symposium

[3] Dalia Ali, Dora Zhao, Allison Koenecke, and Orestis Papakyriakopoulos. 2025. Operationalizing Pluralistic Values in Large Language Model Alignment Reveals Trade-offs in Safety, Inclusivity, and Model Behavior. *arXiv preprint arXiv:2511.14476* (2025).

[4] American Council on Education. 2025. Carnegie Classification of Institutions of Higher Education®. https://carnegieclassifications.acenet.edu/.

[5] Mike Ananny and Kate Crawford. 2018. Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *new media & society* 20, 3 (2018), 973–989.

[6] Clare Baek, Tamara Tate, and Mark Warschauer. 2024. "ChatGPT seems too good to be true": College students' use and perceptions of generative AI. *Computers and Education: Artificial Intelligence* 7 (2024), 100294.

[7] Dean W. Ball and Ketan Ramakrishnan. 2025. Entity-Based Regulation in Frontier AI Governance. *Carnegie Endowment for International Peace* (2025). https://carnegieendowment.org/research/2025/06/artificial-intelligence-regulation-united-states?lang=en Published July 7, 2025.

[8] William Blackstone. 1765. *Commentaries on the Laws of England.* Clarendon Press, Oxford. Book 4, Chapter 27 has the quote..

[9] Center for Innovative Teaching and Learning. 2024. AI Detectors: An Ethical Minefield. *NIU Center for Innovative Teaching and Learning News* (2024). https://citl.news.niu.edu/2024/12/12/ai-detectors-an-ethical-minefield/

[10] Cecilia Ka Yuk Chan. 2023. A comprehensive AI policy education framework for university teaching and learning. *International journal of educational technology in higher education* 20, 1 (2023), 38.

[11] Yucheng Chu, Hang Li, Kaiqi Yang, Harry Shomer, Yasemin Copur-Gencturk, Leonora Kaldaras, Kevin Haudek, Joseph Krajcik, Namsoo Shin, Hui Liu, and Jiliang Tang. 2025. A LLM-Powered Automatic Grading Framework with Human-Level Guidelines Optimization. In *Proceedings of the 18th International Conference on Educational Data Mining*, Caitlin Mills, Giora Alexandron, Davide Taibi, Giosuè Lo Bosco, and Luc Paquette (Eds.). International Educational Data Mining Society, Palermo, Italy, 31–41. doi:10.5281/zenodo.15870201

[12] Simon Coghlan, Tim Miller, and Jeannie Paterson. 2021. Good proctor or "big brother"? Ethics of online exam supervision technologies. *Philosophy & Technology* 34, 4 (2021), 1581–1606.

[13] Glenn Colby. 2025. Data Snapshot: Tenure and Contingency in US Higher Education, Fall 2023. *Academe* 111, 2 (Spring 2025). https://www.aaup.org/academe/issues/spring-2025/data-snapshot-tenure-and-contingency-us-higher-education-fall-2023

[14] Rianne Conijn, Ad Kleingeld, Uwe Matzat, and Chris Snijders. 2022. The fear of big brother: The potential negative side-effects of proctored exams. *Journal of Computer Assisted Learning* 38, 6 (2022), 1521–1534.

[15] Loredana Fattorini, Nestor Maslej, Raymond Perrault, Vanessa Parli, John Etchemendy, Yoav Shoham, and Katrina Ligett. 2024. *The Global AI Vibrancy Tool.* Technical Report. AI Index Project, Institute for Human-Centered AI, Stanford University. https://hai.stanford.edu/assets/files/global_ai_vibrancy_tool_paper_november2024.pdf November 2024 edition.

[16] Rafael Ferreira Mello, Cleon Pereira Junior, Luiz Rodrigues, Filipe Dwan Pereira, Luciano Cabral, Newarney Costa, Geber Ramalho, and Dragan Gasevic. 2025. Automatic Short Answer Grading in the LLM Era: Does GPT-4 with Prompt Engineering beat Traditional Models?. In *Proceedings of the 15th International Learning Analytics and Knowledge Conference (LAK '25).* Association for Computing Machinery, New York, NY, USA, 93–103. doi:10.1145/3706468.3706481

[17] University Center for Teaching and University of Pittsburgh Learning. 2025. Encouraging Academic Integrity. Online resource. https://teaching.pitt.edu/resources/encouraging-academic-integrity/

[18] Megan Fritts. 2025. A Matter of Words: What can university AI committees actually do? *The Point Magazine – Examined Life* (12 May 2025). https://thepointmag.com/examined-life/a-matter-of-words/ Accessed: 2025-10-01.

[19] Lon Luvois Fuller. 1964. *The Morality of Law.* Yale University Press, New Haven, CT.

[20] Andrés Herrera-Poyatos, Javier Del Ser, Marcos López de Prado, Fei-Yue Wang, Enrique Herrera-Viedma, and Francisco Herrera. 2025. Responsible Artificial Intelligence Systems: A Roadmap to Society's Trust through Trustworthy AI, Auditability, Accountability, and Governance. arXiv:2503.04739 [cs.CY] https://arxiv.org/abs/2503.04739

[21] Kalley Huang. 2023. Alarmed by AI chatbots, universities start revamping how they teach. *The New York Times* 16 (2023), e12255.

[22] Panos Ipeirotis. 2025. *Fighting Fire with Fire: Scalable Personalized Oral Exams with an ElevenLabs Voice AI Agent.* https://www.behind-the-enemy-lines.com/2025/12/fighting-fire-with-fire-scalable-oral.html

[23] Sheila Jasanoff et al. 2004. *States of knowledge.* Taylor & Francis Abingdon, UK.

[24] Michael Katell, Meg Young, Dharma Dailey, Bernease Herman, Vivian Guetler, Aaron Tam, Corinne Bintz, Daniella Raz, and P. M. Krafft. 2020. Toward situated interventions for algorithmic equity: lessons from the field. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (Barcelona, Spain) *(FAT* '20).* Association for Computing Machinery, New York, NY, USA, 45–55. doi:10.1145/3351095.3372874

[25] Waleed Khattak. 2025. Algorithmic Due Process Audits: A Legal Framework for AI Accountability in Government Decisionmaking. *Available at SSRN 5362310* (2025).

[26] Niamh Kinchin. 2024. "Voiceless": the procedural gap in algorithmic justice. *International Journal of Law and Information Technology* 32 (2024), eaae024.

[27] Weixin Liang, Mert Yuksekgonul, Yining Mao, Eric Wu, and James Zou. 2023. GPT detectors are biased against non-native English writers. *Patterns* 4, 7 (2023), 100779. doi:10.1016/j.patter.2023.100779

[28] Stefanie A Lindquist and Frank C Cross. 2010. Stability, predictability and the rule of law: stare decisis as reciprocity norm. In *Conference on Legal Decision Making (Indiana)*.

[29] Henrietta Lyons, Eduardo Velloso, and Tim Miller. 2021. Conceptualising contestability: Perspectives on contesting algorithmic decisions. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–25.

[30] Florence Martin, Drew Polly, Shanna Coles, and Chuang Wang. 2020. Examining higher education faculty use of current digital technologies: Importance, competence, and motivation. *International Journal of Teaching and Learning in Higher Education* 32, 1 (2020), 73–86.

[31] Nora McDonald, Aditya Johri, Areej Ali, and Aayushi Hingle Collier. 2025. Generative artificial intelligence in higher education: Evidence from an analysis of institutional policies and guidelines. *Computers in Human Behavior: Artificial Humans* 3 (2025), 100121.

[32] Tortoise Media. 2024. The Global Artificial Intelligence Index: Methodology Report. https://www.tortoisemedia.com/_app/immutable/assets/AI-Methodology-2409.BGTLUPC-.pdf

[33] Jacob Metcalf, Emanuel Moss, Elizabeth Anne Watkins, Ranjit Singh, and Madeleine Clare Elish. 2021. Algorithmic impact assessments and accountability: The co-construction of impacts. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*. 735–746.

[34] Alexi Mostrous, Serena Cesareo, and Joe White. 2024. The Global Artificial Intelligence Index 2024. *Tortoise Media* (Sep 2024). https://www.tortoisemedia.com/2024/09/18/the-global-artificial-intelligence-index-2024

[35] Matti Mäntymäki, Matti Minkkinen, Teemu Birkstedt, and Mika Viljanen. 2023. Putting AI Ethics into Practice: The Hourglass Model of Organizational AI Governance. arXiv:2206.00335 [cs.AI] https://arxiv.org/abs/2206.00335

[36] National Endowment for the Humanities. 2025. Humanities Perspectives on Artificial Intelligence. https://www.neh.gov/AI.

[37] National Institute of Standards and Technology. 2023. *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. NIST Special Publication. National Institute of Standards and Technology (NIST). https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf

[38] Pablo Fuentes Nettel, Emma Hankins, Richard Stirling, Giulia Cirri, Gonzalo Grau, Sulamaan Rahim, and Eddie Crampton. 2024. *Government AI Readiness Index 2024*. Report 7th Edition. Oxford Insights.

[39] Kavitha Palaniappan, Elaine Yan Ting Lin, and Silke Vogel. 2024. Global regulatory frameworks for the use of artificial intelligence (AI) in the healthcare services sector. In *Healthcare*, Vol. 12. MDPI, 562.

[40] Emmanouil Papagiannidis, Patrick Mikalef, and Kieran Conboy. 2025. Responsible artificial intelligence governance: A review and research framework. *The Journal of Strategic Information Systems* 34, 2 (2025), 101885.

[41] Britt S. Paris, Cynthia Conti-Cook, Daniel Greene, Kyle M. L. Jones, Brian Justie, Matthew Kirschenbaum, Lisa Kresge, Emma May, Aiha Nguyen, Rebecca Reynolds, et al. 2025. Artificial Intelligence and Academic Professions. *Academe* 111, 3 (July 2025), 49–59.

[42] Jessica Priest. 2025. Texas universities deploy AI tools to review and rewrite how some courses discuss race and gender. *The Texas Tribune* (15 Dec. 2025). https://www.texastribune.org/2025/12/15/texas-universities-ai-course-audits/

[43] Marie T Reilly. 2015. Due process in public university discipline cases. *Penn St. L. Rev.* 120 (2015), 1001.

[44] Respondus, Inc. 2026. *Respondus*. Respondus, Inc. https://web.respondus.com/

[45] Teresa Scassa. 2022. The surveillant university: Remote proctoring, AI, and human rights. *Can. J. Comp. & Contemp. L.* 8 (2022), 271.

[46] Andrew D Selbst. 2021. An institutional view of algorithmic impact assessments. *Harv. JL & Tech.* 35 (2021), 117.

[47] Andrew D. Selbst, Danah Boyd, Sorelle A. Friedler, Suresh Venkatasubramanian, and Janet Vertesi. 2019. Fairness and Abstraction in Sociotechnical Systems. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (Atlanta, GA, USA) *(FAT* '19)*. Association for Computing Machinery, New York, NY, USA, 59–68. doi:10.1145/3287560.3287598

[48] Oleksan V Spivakovsky, Serhii A Omelchuk, Vitaliy V Kobets, Nataliia V Valko, and Daria S Malchykova. 2023. Institutional policies on artificial intelligence in university learning, teaching and research. *Information Technologies and Learning Tools* 97, 5 (2023), 181.

[49] Lucy A. Suchman. 2007. *Human-Machine Reconfigurations: Plans and Situated Actions* (2nd ed.). Cambridge University Press, Cambridge, UK.

[50] Araz Taeihagh. 2021. Governance of artificial intelligence. *Policy and society* 40, 2 (2021), 137–157.

[51] Turnitin. 2023. Understanding false positives within our AI writing detection capabilities. https://www.turnitin.com/blog/understanding-false-positives-within-our-ai-writing-detection-capabilities.

[52] United Nations System Chief Executives Board for Coordination. 2024. *UN System White Paper on AI Governance*. Technical Report. UN System Chief Executives Board for Coordination. https://unsceb.org/sites/default/files/2024-11/UNSystemWhitePaperAIGovernance.pdf White Paper, November 2024.

[53] U.S. Census Bureau. n.d.. Census Regions and Divisions of the United States. https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf.

[54] U.S. Department of Education. 2026. U.S. Department of Education Announces Release of $169 Million Under the Fund for the Improvement of Postsecondary Education. https://www.ed.gov/about/news/press-release/us-department-of-education-announces-release-of-169-million-under-fund-improvement-of-postsecondary-education.

[55] James D. Walsh. 2025. Everyone Is Cheating Their Way Through College: ChatGPT has unraveled the entire academic project. *Intelligencer (New York Magazine)* (7 May 2025). https://nymag.com/intelligencer/article/openai-chatgpt-ai-cheating-education-college-students-school.html

[56] Hui Wang, Anh Dang, Zihao Wu, and Son Mac. 2024. Generative AI in higher education: Seeing ChatGPT through universities' policies, resources, and guidelines. *Computers and Education: Artificial Intelligence* 7 (2024), 100326.

University or Organization	Resource Name	Comments	Link
University of La Verne	List of institutions with AI guidelines	*"We found that most universities that we surveyed have some kind of statement or set of guidelines for genAI in the classroom. This list is far from exhaustive."*	Link
Western University of Health Sciences	University Policies on Generative AI	*"Collection of university policies and websites. Questions? Contact CETL@westernu.edu."*	Link
Northeastern University	A moderated list of AI syllabus statements	*"If you would like to submit your course guidelines/policy or revise your submission, please submit it in this form."*	Link
Northeastern University	A moderated list of AI institutional policies	*"This document is maintained by Lance Eaton. You are welcome to share it with other individuals, groups, and organizations. To view the policies, please select the "Policies" tab in this spreadsheet. If you would like to submit your policy, please complete this form (https://bit.ly/AI-Institutional-Policies) and it will show up here within 24-48 hours."*	Link
Gradpilot	The State of AI in College Admissions	*"Navigate AI usage rules across 150+ American universities"*	Link

Table 3. **Prior datasets and resource collections related to institutional AI governance and policy articulation**, situating ACAI-US79 within the broader landscape of governance-focused audits.

[57] Langdon Winner. 1980. Do Artifacts Have Politics? *Daedalus* (1980), 121–136.
[58] Chuhao Wu, He Zhang, and John M. Carroll. 2024. AI Governance in Higher Education: Case Studies of Guidance at Big Ten Universities. arXiv:2409.02017 [cs.HC] https://arxiv.org/abs/2409.02017
[59] Junchao Wu, Shu Yang, Runzhe Zhan, Yulin Yuan, Lidia Sam Chao, and Derek Fai Wong. 2025. A survey on llm-generated text detection: Necessity, methods, and future directions. *Computational Linguistics* 51, 1 (2025), 275–338.
[60] Jiawei Zhang and Boris Paal. 2025. A Three-Layered Framework: An AI Governance Guide for Global Policymakers. (2025). https://papers.ssrn.com/abstract=5241351

A Scope & Limitations

First, our sample is limited to 79 universities, which, although diverse in size and mission, cannot capture the full range of institutional practices globally. Second, the analysis is predominantly USA-centric, reflecting the regulatory, cultural, and policy context of U.S. higher education. Third, our reliance on publicly available institutional data may omit informal practices or internal decision-making processes that shape outcomes but are not externally visible.

B ACAI-US79 Comparable Datasets

As shown in Table 3, several prior efforts examine how universities articulate institutional responses to AI, primarily through collections of publicly available policies, guidelines, and administrative resources, motivating the need for a systematic, governance-focused audit such as ACAI-US79.

C ACAI-US79 Iterative Link Categorization Process

This section describes the iterative link categorization process used to construct ACAI-US79: Figure 5 gives an overview of the process, and Figure 6 shows the detailed annotator instructions for Phases II and III. Because institutional AI governance is unevenly distributed and inconsistently labeled across universities, we employ a recall-oriented, multi-phase procedure that surfaces publicly legible, institutionally authoritative materials while enforcing a clear boundary between governance capacity and general AI research or outreach.

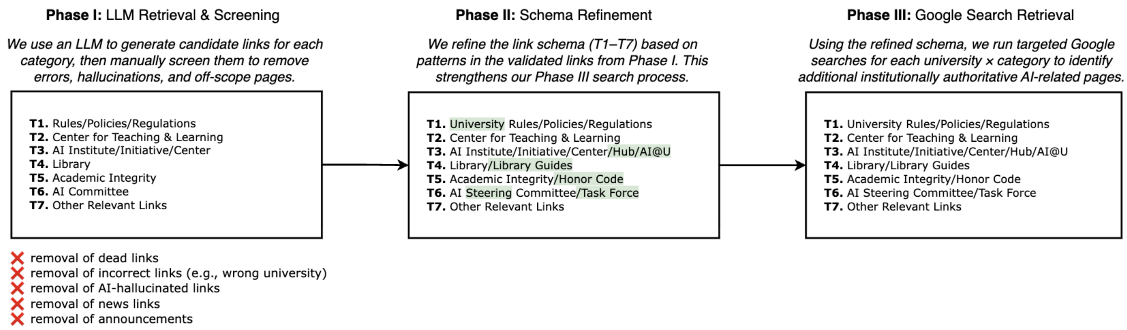


Fig. 5. **Overview of Iterative Link Categorization Process**: See Figure 6 for more details on Phases II and III.

The seven categories (T1–T7) are neither mutually exclusive nor collectively exhaustive. In practice, AI governance materials are unevenly distributed across institutional units or consolidated into a small number of centralized resources. As a result, a single link may be relevant to multiple categories, while other categories may contain no links for a given institution. Such empty cells reflect genuine variation in how AI governance is organized and communicated, not missing data or annotation error. Category membership should therefore be interpreted as evidence of where—and how legibly—governance functions are articulated in public-facing materials.

Phases I and II are deliberately recall-oriented and surface a wide range of AI-related materials, including research and outreach content. **Phase III** enforces the conceptual boundary of the audit by removing links that are AI-related but governance-irrelevant, such as research labs, grants, or faculty-led initiatives that do not articulate institutional authority or procedural expectations. This consolidation step prevents research-intensive institutions from appearing more "governed" simply due to higher volumes of AI-related content and ensures that institutional capacity is not conflated with research productivity. The resulting link set retains only publicly legible, institutionally authoritative governance artifacts, such as standing policies, centrally maintained guidance, and formal committees. By narrowing annotator attention to governance-relevant materials, Phase III also improves annotation consistency and strengthens the construct validity of the audit.

We do not include the following types of links, because we found in Phases I and II of our iterative link categorization process that these types of links generally weren't relevant to our annotation criteria (*) or were separate from our focus in this work (+):

- **Legal / General Counsel***: Especially for risk, compliance, data use, copyright, FERPA, and contracts involving AI tools. We also think that students have access to a variety of AI tools separate from those visible by the university.
- **Research Office / Office of Sponsored Programs**+: Research oversight is mediated through specialized mechanisms (e.g., AI use in grants, data management plans, human subjects, and responsible research conduct) separate from the present audit, which focuses on governance as enacted through core university functions shaping teaching, assessment, and student experience.
- **Data Governance / Privacy Office***: Sometimes separate from IT; increasingly relevant for AI training data and student data use.
- **Accessibility / Disability Services**+: AI accommodations, assistive tech, and equity considerations. While critically important for AI equity, guidance in this area is often individualized/case-specific.

Phase II: LLM Cleanup

For **each university and each category (T1–T7)**, apply a conservative cleanup pass, removing links **only when they are clearly out of scope**. When determining relevance would require nontrivial investigation beyond initial screening, we conservatively retain the link and defer judgment to the annotation process, where deeper inspection is already required. This typically includes links that:

- Are non-authoritative (i.e., not published by the university)
- Are dead, redirecting, or hallucinated
- Are narrowly subject, department, or graduate school specific rather than university-level
- Are strictly research-focused
- Consist solely of news, announcements, or event listings
- Provide application instructions for prospective students
- Are intended for university communications or marketing staff
- Require institutional login for access

Additionally, when verification is straightforward:

- Remove malformed entries (e.g., markdown artifacts instead of valid links)
- Remove excessive sublinks pointing to the same underlying page
- De-duplicate or re-categorize links

Phase III: Link Identification

For **each university and each category (T1–T6)**, if a clear main link has not already been identified:

(1) **Identify the Authoritative Institutional Surface and Any Clearly AI-Related Subpages** (Skip to Step #2 for T1)
  (a) Perform a Google search for "University" + "Category"
  (b) Navigate to the primary institutional landing page(s)
    (i) Review standard navigation paths (e.g., scrolling the page, examining menu bars)
    (ii) Include any subpages that are clearly AI-related
    (iii) If additional clearly AI-related subpages are encountered during exploration, include them
(2) **Targeted Confirmation Search** (If no AI-related content is visible in Step #1; except T3 and T6, which explicitly reference "AI")
  (a) Perform a Google search for "University" + "Category" + "AI"
  (b) Include any institutionally authoritative pages that are clearly relevant
(3) If no suitable links are identified, **record "None" for that category.**
(4) Mark the university as complete once all categories are reviewed.

Fig. 6. **Annotation Instructions for Phases II and III of Iterative Link Categorization Process.**

- **Human Resources$^+$**: Staff and faculty use of AI for hiring, evaluation, or administrative work.
- **Admissions$^+$**: Policies on AI-assisted application materials are often separate from academic integrity rules.
- **Graduate School$^+$**: Graduate-specific guidance often differs from undergraduate rules.
- **Department or College-Level Pages$^+$**: Many institutions defer AI guidance to colleges (e.g., Engineering, Business, Law) or even individual departments.

This dataset reflects publicly available, institutionally maintained web resources and therefore has a few key limitations. First, universities differ substantially in how AI-related guidance is organized, labeled, and distributed across administrative units, which may lead to uneven coverage across institutions. We aimed to yield equitable coverage with our iterative link categorization process (Figure 5) as to correctly capture the heterogeneous ways in which different universities – varying in region, research activity, student population size, and more as discussed in §?? – share and present AI policy. Second, the dataset captures only formal, publicly visible governance artifacts and does not reflect informal practices, internal guidance, or unpublished decision-making processes that may substantially shape how AI is used and regulated within institutions. As a result, the dataset should be interpreted as a representation of officially articulated AI governance rather than a complete account of institutional practice.

C.1 Phase II vs. III Coverage Comparison

In this section, we provide details on the links retrieved in Phase II vs. III of our iterative link categorization process, characterizing how human annotators empowered by Google Search in Phase III processed the LLM output results from Phase II. Specifically, for each institution-category pair, we compare the LLM-retrieved link set L with the human-curated set H using set-based metrics. We compute precision ($|H \cap L|/|L|$) and recall ($|H \cap L|/|H|$), along with the number of *added* links ($|L \setminus H|$) and *deleted* links ($|H \setminus L|$). We interpret recall as a measure of *coverage*, indicating whether the LLM retrieves policy evidence aligned with human judgment.

How good is LLM coverage of institutional AI policy? In Figure 7a, we show recall (human coverage) against precision (LLM correctness) for all institution-category pairs. The distribution shows that LLM retrieval is strongly recall-oriented: many cases achieve high recall but only moderate precision, indicating that while the LLM often retrieves at least one relevant policy link, it frequently includes additional links that are later removed by human annotators. Instances of low recall correspond to true coverage failures, where the LLM fails to retrieve any human-recognized policy evidence. To further quantify this behavior at the category level, we summarize LLM coverage using three increasingly strict definitions: *raw coverage* (the presence of any LLM-retrieved link), *aligned coverage* (overlap with human-curated links), and *missed coverage* (human-curated links not retrieved by the LLM). The results are shown in Figure 7b. While raw coverage is consistently high across categories, aligned coverage drops substantially, indicating that apparent LLM coverage overstates effective policy coverage. This discrepancy is most pronounced in categories involving decentralized or evolving institutional resources, such as libraries and AI initiatives, where authoritative policy evidence is more difficult to identify automatically. Taken together, these results suggest that LLMs perform well as a first-stage retrieval mechanism, successfully surfacing candidate AI policy evidence for most institutions and categories. However, the observed gap between apparent and aligned coverage underscores the continued necessity of human supervision to filter non-authoritative links, recover missed policy documents, and ensure that final policy representations accurately reflect institutional governance.

C.2 Example Links

In Table 4, we provide examples of the types of links retrieved in T1-T7.

Table 4. **Examples of Pages by Link Type (T1–T7):** This table provides illustrative examples of pages classified under each link type, chosen to help readers understand how link categories are defined and applied in the search framework.

University	Quote	Link
	T1. University Rules/Policies/Regulations	
University of Texas at Austin	With the increasing integration of artificial intelligence (AI) tools—such as ChatGPT, Copilot, Bard Gemini, Claude and other generative AI applications known as large language models (LLM), diffusion models, or generative AI applications—into university activities, it's essential to use these technologies responsibly. This guidance, developed collaboratively by the Office of Legal Affairs, University Compliance Services, the Information Security Office, and the Business Contracts Office, outlines acceptable practices for utilizing generative AI tools while safeguarding institutional, personal, and proprietary information. Additional guidance may be forthcoming as circumstances evolve.	Link

Continued on next page

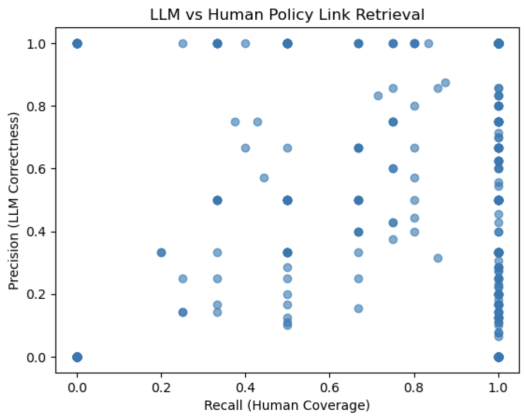
University	Quote	Link
Georgia Southern University	This policy establishes guidance for the responsible, ethical, and transparent use of Artificial Intelligence (AI) tools at Georgia Southern University (University) including in teaching, learning, assessment, classroom activities, University community service, research, creative activity, scholarly communication, and administrative activities, while encouraging innovation, academic freedom, and appropriate autonomy while maintaining compliance with all data security and privacy regulations. This policy ensures that the University complies with Board of Regents (BOR) Policy 6.28 Artificial Intelligence in Academic Context; federal, state, and international laws; and industry standards and best practices.	Link
Carleton College	AI technologies raise novel questions around data security, attribution, and ethics. In many cases, Carleton's existing policies still apply to AI, but in some cases this technology requires new policies or new interpretations of existing policies. This page will provide links to policies that apply directly to AI use at Carleton and highlight any additions and changes as they're made.	Link
University of Southern California	This Research Guide provides information on the use of Generative AI in academic papers and research, and provides guidance on the ethical use of Generative AI in an academic setting.	Link
	T2. Center for Teaching & Learning	
Stanford University	AI Meets Education at Stanford (AIMES) is a VPUE effort to catalyze and support critical engagement with generative AI in Stanford teaching and learning contexts, coordinated by the Center for Teaching and Learning.	Link
Ball State University	Explore a variety of courses designed to meet learners at all levels. Whether you're looking for introductory classes or advanced specialization tracks, these courses provide structured, in-depth instruction in AI topics to build your skills and confidence. A strong grasp of AI terminology is essential for navigating complex concepts and discussions. This glossary of terms offers definitions and explanations of key terms, serving as a quick reference to clarify AI language as you advance in your learning.	Link
University of Michigan at Ann Arbor	The release of ChatGPT in late 2022 jump started an ongoing and growing exchange in higher education about both the promises and significant risks posed by Generative Artificial Intelligence, particularly to the teaching and learning enterprise. This site is designed to offer links to programs and resources from U-M and beyond to help you navigate this new landscape. Given how rapidly the GenAI landscape is shifting, we include links to sources that offer regular posts and updates on this topic.	Link
University of Florida	How is Artificial Intelligence (AI) affecting teaching and learning in higher education? Artificial intelligence (AI) is significantly impacting higher education, revolutionizing various aspects of the learning experience. AI-powered tools and platforms are transforming how students access educational content, tailor their learning paths, and receive personalized feedback. Moreover, AI-driven systems can help educators generate educational content and facilitate research endeavors. While these advancements bring exciting opportunities, it's essential to address ethical concerns, data privacy, and ensure AI complements the education process rather than replacing it entirely. Embracing AI responsibly can lead to a more accessible, efficient, and effective higher education landscape. Artificial intelligence has introduced significant challenges to academic integrity in education. As AI becomes more accessible, educators have expressed concerns about students using it to generate answers to questions on tests and assignments. Rather than reacting in fearful ways to new advances in AI, educators can focus on potential benefits, such as providing new perspectives on a problem and generating content that can be analyzed or critiqued. Undoubtedly, faculty need to provide guidelines to students about the appropriate and inappropriate uses of AI tools. However, faculty can also model and encourage productive and positive uses of AI and help students see its value.	Link
	T3. AI Institute/Initiative/Center/Hub/AI@U	
University of Wyoming	The University of Wyoming's AI Initiative is a bold, people-centered effort to shape the future of our state, empowering citizens and communities to thrive in an AI-driven world. By addressing key industries like agriculture, engineering, energy, tourism, wildlife conservation, and rural healthcare, UW is ensuring that AI enriches lives and drives sustainable growth. This initiative will enhance the University of Wyoming's ability to bring advances in AI to disciplines across the university to advance the state. It will attract investments, build corporate partnerships, seed entrepreneurship, and equip every student and community to participate in the global AI transformation, securing a prosperous future for all of Wyoming.	Link

Continued on next page

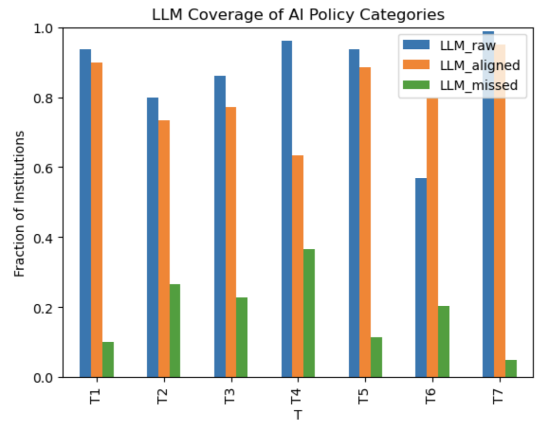
University	Quote	Link
University of California, Berkeley	Welcome to the AI Hub – your central resource for artificial intelligence at UC Berkeley. Rooted in Berkeley's pioneering spirit, ethos of inclusivity, and culture of excellence, this hub connects our community with essential AI tools, training, policies, and opportunities. Whether you're a student, researcher, faculty, or staff member, you'll find guidance, collaboration, and innovation here to help navigate the evolving world of AI and amplify the impact of our collective efforts.	Link
San José State University	Welcome to the bold new world of Artificial Intelligence (AI), where groundbreaking innovation meets inclusive leadership in the heart of Silicon Valley. At San José State University, we are a place of firsts, pioneering advancements in AI and empowering the next generation of leaders to shape the future of technology. Our interdisciplinary programs blend cutting-edge research with hands-on learning, equipping students and professionals to solve real-world challenges and explore ethical solutions in AI. As a proud partner in Silicon Valley's ecosystem of global innovation, SJSU connects you to industry leaders, transformational opportunities, and a vibrant community. Whether you're forging new paths in AI development, exploring its societal impact or preparing to lead in this dynamic field we'll help you unlock your potential and create a future where everyone can thrive. San José State University leads the way in AI innovation and leadership.	Link
Clark Atlanta University	The NSF Expand-AI project led by Clark Atlanta University (CAU) in collaboration with AI4OPT builds an AI Hub at CAU to transform accessibility to AI jobs, AI research, and the AI ecosystem. AIHub@CAU consists of three key pillars: (1) a Master Program in AI; (2) a PhD program in AI; and (3) research collaborations between CAU and Georgia Tech. The program is a joint project by the department of mathematical sciences, the department of Cyber-Physical Systems and the School of Business Administration, making it truly multidisciplinary.	Link
	T4. Library/Library Guides	
Mercer University	Definitions. Artificial Intelligence: AI is typically defined as the ability of a machine to perform cognitive functions we associate with human minds, such as perceiving, reasoning, learning, and problem solving. Examples of technologies that enable AI to solve complex problems include robotics, computer vision, language, virtual agents and machine learning.	Link
Chapman University	AI Literacy. Use this guide to understand Artificial Intelligence literacy in the context of higher education. What is AI Literacy? "AI literacy is the ability to understand, use, and think critically about AI technologies and their impact on society, ethics, and everyday life." - Lo, L. S. (2025). AI Literacy: A Guide for Academic Libraries. College & Research Libraries News, 86(3), Article 3. https://doi.org/10.5860/crln.86.3.120	Link
University of Chicago	Generative AI. Information about Generative AI tools and their use in and outside of the classroom.	Link
Brown University	Generative artificial intelligence has already started to have an impact on the way we discover, manage, create, and disseminate information. Generative AI tools are in a state of rapid development, and new information about applications, policies, and social impact is released each day. While every attempt will be made to keep this guide up to date, please be aware that the information included here is likely to age quickly. This guide includes context and advice for engaging with generative Artificial Intelligence, and does not represent University policy. The Library does not endorse any specific AI technologies, and encourages users to be cautious about sharing personal information when using AI tools.	Link
	T5. Academic Integrity/Honor Code	
Creighton University	Cheating: The deliberate use or attempted use of unauthorized material in an academic exercise, including unauthorized collaboration with classmates, or use of unauthorized work created by artificial intelligence.	Link
Wofford College	Unauthorized use of generative artificial intelligence to create content that is submitted as one's own.	Link
Marquette University	Academic Integrity. A Message for Faculty, Staff, and Students on the use of Large Language Model-Based Chatbots ("generative artificial intelligence") at Marquette University.	Link
University of New Hampshire	Cheating. Use or attempted use of any academic exercise materials, information, study aids, electronic data, AI tools, assignment/exam surrogate, or other forms of assistance without authorization.	Link
	T6. AI Steering Committee/Task Force	

Continued on next page

University	Quote	Link
Texas A&M University	The Artificial Intelligence, Innovative & Emerging Technologies Work Group equips faculty with the knowledge and resources to explore and integrate cutting-edge technologies into teaching and learning. By curating and sharing AI-related resources, fostering collaboration, and promoting best practices, the group empowers educators to leverage emerging technologies to enhance student engagement and academic success across the A&M System.	Link
Stony Brook University	Library AI Steering Committee. The Stony Brook University Libraries AI Steering Committee plays a central role in guiding the responsible and strategic integration of artificial intelligence across library services, operations, and research support. Established to ensure that emerging technologies advance—rather than compromise—the Libraries' core values of equity, accessibility, intellectual freedom, and responsible innovation, the committee evaluates opportunities and risks, recommends best practices, and supports evidence-based decision-making. Its charge includes reviewing AI initiatives for alignment with institutional priorities, developing ethical guidelines and principles, promoting staff training and AI literacy, and fostering collaborations with campus partners and professional communities. Through regular reporting and transparent communication with library leadership, the AI Steering Committee helps ensure accountability and positions the Libraries to thoughtfully and proactively navigate the evolving landscape of AI in higher education.	Link
California State University, Long Beach	The purpose of the AI Academic Subcommittee is to explore AI technologies and plan for future implementations. The subcommittee will make recommendations to the AI Steering Committee to develop guidelines for campus-wide deployment. To foster a community of AI users on campus, the subcommittee will also make recommendations for professional development and support for faculty and staff on AI-related topics.	Link
Iowa State University	The 2024 Generative AI Guidance Committee has successfully completed its charge. The subcommittees were assembled, carried out their tasks diligently, and contributed valuable insights and recommendations. Their work has laid a strong foundation for our institutional AI strategy. For more information and access to resources, please visit ai.iastate.edu.	Link
	T7. Other Relevant Links	
University of Michigan at Ann Arbor	Custom GenAI Services for the U-M Community. U-M is proud to be the first university in the world to provide a custom suite of generative AI tools to its community. With a focus on equity, accessibility, and privacy, our AI Services are available to all U-M faculty, staff, and students on the Ann Arbor, Flint, Dearborn, and Michigan Medicine campuses.	Link
University of Chicago	Advances in artificial intelligence (AI) and data science are driving breakthroughs – transforming scientific discovery, accelerating innovation, and changing entire industries. At the University of Chicago, our long-standing tradition of rigorous inquiry and interdisciplinary collaboration provides a powerful foundation for tackling large-scale problems in AI and data science and unleashing their greatest potential to improve individual lives and our world. Faculty and researchers from every division and school at UChicago are at the forefront of these fields, from developing trustworthy AI systems and foundational frameworks to enabling transformative advances in areas such as precision medicine and next-generation climate modeling.	Link
Northern Illinois University	NIU has introduced Mission, a new feature for undergraduate students, that includes reaching out to you through AI-assisted text messages and a chatbot to respond to your questions around the clock. Mission will help answer your questions, and connect you to campus services and information, including: Academic success - tutors, academic advisors, study skills and more. Financial matters - financial aid, FAFSA and more. Student life and involvement - clubs, organizations, events and more. Well-being/mental and physical health - counseling services, nutrition, campus recreation and more. Mission will provide timely, accurate responses via text to your questions at all times of day, regardless of your location. No logins or app downloads are required.	Link
Case Western Reserve University	University Technology offers many services and applications related to Generative AI. Below are some AI technologies available to the campus community. Note: Consumer AI services, especially free ones, often collect the data you enter into them and use that data in their training models. This can lead to your data being made available via these AI services. Never put sensitive university information into an AI service if the university does not have a contract with the AI vendor with proper privacy and security safeguards. The university offers AI services that will protect your data. Use those services when sensitive information is involved.	Link



(a) Precision–recall characteristics of LLM-based AI policy link retrieval (Phase II) compared to human-curated links (Phase III). Each point represents one institution–category pair.



(b) LLM coverage of institutional AI policies under three definitions: raw coverage, aligned coverage, and missed coverage.

Fig. 7. Comparison of LLM-based and human-curated AI policy link retrieval.

C.3 Prolific Privacy Policy

We provide the Privacy Policy in Figure 8.

D ACAI Rank Changes Under Alternate Weighting Schemes

To assess whether ACAI rankings are artifacts of normative weighting choices, we evaluated four weighting schemes: an indicator-weighted baseline ($A = 2, B = 4, C = 3, D = 2$), equal ($A = 1, B = 1, C = 1, D = 1$), policy-heavy ($A = 1, B = 1, C = 2, D = 2$), and teaching-heavy ($A = 1, B = 2, C = 1, D = 1$), using percentile ranks. We show the results in Tables 5 and 6.

Table 5. Pairwise Rank Correlations Across Schemes

Scheme 1	Scheme 2	Spearman	Pearson
Baseline	Equal	0.98	0.98
Baseline	Policy-heavy	0.93	0.93
Baseline	Teaching-heavy	0.99	0.99
Equal	Policy-heavy	0.97	0.97
Equal	Teaching-heavy	0.98	0.98
Policy-heavy	Teaching-heavy	0.93	0.93

E Details on Comparison to CSRankings

To contextualize institutional AI governance capacity relative to AI research activity, we constructed a reference list of research-active universities using CSRankings, shown in Figure 9, a widely used, publicly available ranking of computer science research output. CSRankings aggregates publication counts across major computer science venues and allows filtering by research area and time period.

Privacy Policy
Last Updated: December 29, 2025
Platform: Prolific (https://www.prolific.com)
If you have questions about this study, you may contact the research team at `redacted@redacted`.

1. Purpose of the Task
You are being asked to participate as an **annotator** in a research project examining how U.S. universities define and respond to **artificial intelligence (AI)** in their institutional policies and teaching resources. Your role involves reviewing university web pages and coding the presence or absence of certain indicators related to AI governance, academic integrity, and instructional support. The study analyzes publicly available institutional documents and does not evaluate individual instructors, students, or staff.
This task contributes to a larger academic study focused on understanding patterns in higher education responses to AI technologies.

2. Data We Collect
During your participation, we collect the following categories of data:

- **Prolific ID and Prolific-provided Demographic Data:** This data is provided by the Prolific platform to us. Your Prolific ID will be used only for compensation and quality control, and will be removed as part of the anonymization process before dataset release.
- **Timing Metadata:** The start and completion time for your task submission, which helps assess annotation duration and data quality.
- **Annotation Data:** Your coded responses and URLs that you provide.

No personal browsing history, IP address, or system-level data is collected by the researchers; such information remains with Prolific and is governed by their Privacy Policy.

3. How Your Data Is Used
Your coded responses will be:

- **De-identified.**
- **Used for academic research and publication** in peer-reviewed journals or conference presentations.

De-identified datasets will be shared publicly for use by other researchers under ethical data-sharing agreements, consistent with open science practices.

4. Data Storage and Security
All data collected will be stored securely in **encrypted storage** (e.g., Google cloud) accessible only to the research team.

- Data will be **retained for up to 3 years** after study completion and then deleted or permanently anonymized.
- No data will be sold or shared with commercial entities.

5. Voluntary Participation and Withdrawal
Your participation is **entirely voluntary**. You may withdraw from the task at any time prior to submission on Prolific. If you withdraw before completing the task, no partial data will be used.

6. Risks and Benefits
There are **minimal risks** associated with participation. You may experience minor fatigue from reviewing university materials. There are **no direct personal benefits**, though your work contributes to research improving understanding of AI use in education policy.

7. Confidentiality
Your responses will never be linked to your name or contact information. Any publications or presentations resulting from this research will contain only **anonymous findings**.

8. Consent
By completing the task on Prolific, you confirm that you:

- Are 18 years of age or older.
- Understand the nature and purpose of this research.
- Consent to your **anonymized** responses being used for research and publication purposes.

Fig. 8. **Privacy Policy Provided to Human Annotators Recruited via the Prolific Platform.**

We first navigated to csrankings.org and restricted the ranking to AI-relevant research areas only, enabling the categories of Artificial Intelligence, Computer Vision, Machine Learning, Natural Language Processing, and The Web

Table 6. Maximum Absolute Rank Change Statistics

	Baseline	Equal	Policy-heavy	Teaching-heavy
Mean	16.28	17.37	16.96	16.54
Median (50%)	14.50	17.00	15.50	13.50
75%	22.50	23.50	22.50	21.75
Max	46.00	46.00	43.50	45.00

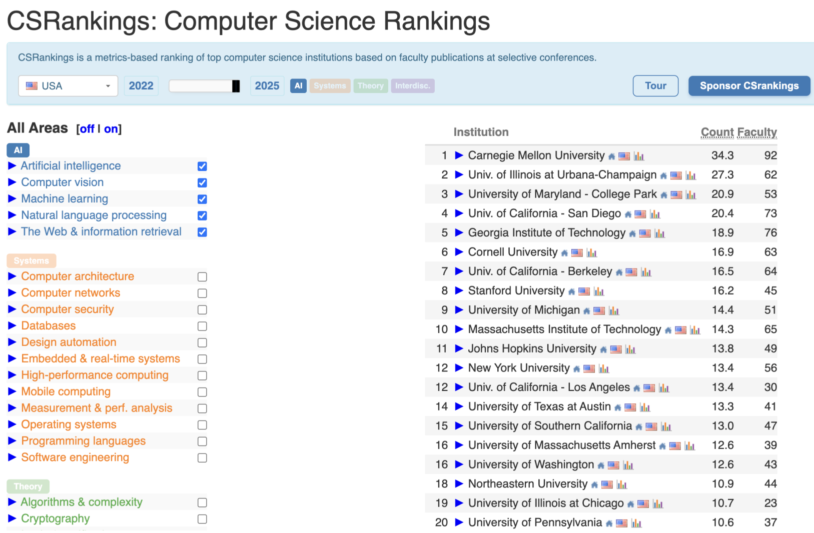


Fig. 9. **CSRankings Website**: Used to approximate research output, https://csrankings.org.

& Information Retrieval, while disabling all other areas. This filtering step was intended to approximate institutional engagement with AI-related research rather than overall computer science output.

We then restricted the publication window to 2022–2025, corresponding to the period following the public release of large-scale generative AI systems (i.e., ChatGPT) and the following rapid expansion of AI use in educational contexts. This temporal filter was chosen to reflect contemporary AI research activity during the period in which universities began articulating institutional responses to generative AI.

The resulting ranked list of universities was used as a reference set for identifying institutions with substantial recent AI research activity. Importantly, this list was not treated as a measure of governance quality or institutional responsibility. Rather, it served as a comparative baseline to examine whether AI research intensity correlates with publicly articulated AI governance capacity.

This procedure relies exclusively on publicly accessible filters and settings within CSRankings and is fully reproducible.

(a) Exclude unranked CSRankings$_{AI}$; results also visualized in Figure 3

Group	n	Pearson	Spearman	Kendall τ	Spearman CI$_{Low}$	Spearman CI$_{High}$
South	9	0.51	0.65	0.42	-0.05	0.93
Midwest	10	0.26	0.38	0.27	-0.35	0.85
West	7	0.25	0.18	0.24	-0.76	1.00
Northeast	9	0.05	0.09	-0.03	-0.72	0.66
R1	29	0.14	0.12	0.08	-0.25	0.49
R2	6	-0.31	-0.32	-0.28	-1.00	0.89
Public Research	18	0.05	-0.05	-0.06	-0.65	0.54
Private Research	17	0.20	0.22	0.17	-0.32	0.61
Medium	9	0.32	0.27	0.17	-0.62	0.83
Large	21	0.08	0.02	0.00	-0.49	0.50

(b) Include unranked CSRankings$_{AI}$ by collapsing into bottom rank

Group	n	Pearson	Spearman	Kendall τ	Spearman CI$_{Low}$	Spearman CI$_{High}$
South	21	0.63	0.70	0.55	0.32	0.87
Midwest	20	0.49	0.54	0.42	0.14	0.80
Northeast	20	0.35	0.40	0.29	-0.04	0.73
West	18	0.35	0.34	0.26	-0.18	0.70
R1	33	0.24	0.24	0.16	-0.11	0.55
R2	30	0.00	0.11	0.08	-0.27	0.47
Public Research	33	0.44	0.47	0.32	0.16	0.71
Private Research	30	0.35	0.41	0.32	0.06	0.71
Small	27	-0.17	-0.04	-0.02	-0.41	0.43
Medium	25	0.45	0.49	0.37	0.09	0.78
Large	27	0.26	0.25	0.16	-0.18	0.61

Table 7. Agreement between ACAI and CSRankings$_{AI}$ across institutional groups with $n > 5$. The table reports Pearson, Spearman, and Kendall correlations. Spearman confidence intervals are based on 1000 bootstrap resamples.

F Extended Related Work

F.1 Institutional & Socio-Technical Systems Theory, and Algorithmic Accountability

As Winner argued, *"What matters is not technology itself, but the social or economic system in which it is embedded"* [57]. This position, described as social determination theory, offers a corrective to naïve technological determinism: the assumption that technology evolves according to its own internal logic and subsequently shapes society in a one-directional way. Technological and institutional capacities do not evolve independently, instead co-producing and mutually constructing each other [23]. This aligns with the argument of Selbst et al. [47], who warn that abstracting away from social context obscures the structural forces that shape fairness itself.

Algorithmic accountability is usually determined via algorithmic impact assessments, an auditing mechanism to judge how an algorithm is causing harm. Metcalf et al. [33] highlights the distinctiveness of algorithmic systems, noting that there is a heightened risk with misunderstanding the inner workings of algorithmic systems, and therefore being unable to legislate their use or development effectively. Selbst [46] also note this gap between the inner workings of a computational system and effective governance, demonstrating cases of "algorithmic harm where existing liability regimes fail to hold the creators of the harm to account, specifically because of a lack of knowledge about the development process." Notably, Ananny and Crawford [5] criticize the transparency ideal in algorithmic assessments, noting that "transparency alone cannot create accountable systems."

F.2 AI Governance Models

Recently, there have been calls for attention to AI governance [39, 40, 50]. Prior AI governance models formulate the relationship between governance and computation as unidirectional capacity flows, as shown in Table 8. Across this literature, governance and computation are treated as sequential processes, where one directs, regulates, or reacts to the other. These frameworks have advanced the field's understanding of risk and ethics, but they remain largely reactive and linear, conceptualizing governance as either a *top-down constraint* or a *downstream response* to technical innovation.

Existing models recognize the importance of institutional capacity: The UN System Survey of Institutional Models [52] emphasizes "capacity-building" but defines it primarily as the technical training of scientists and regulators; The Responsible AI Systems Roadmap [20] focuses on the role of scientists in shaping policy. While models recognize the importance of institutional development, they largely equate capacity with technical skill or regulatory compliance rather than with civic or educational infrastructure.

Model	Primary Focus	Capacity Flow	Scope
The Hourglass Model of Organizational AI Governance [35]	Focused on AI ethics, risk mitigation; organized into environmental, organizational, and AI system layers with some feedback mechanisms (i.e. computational → governance).	Governance → Computational	EU/Global
NIST AI Risk Management Framework [37]	Focused on identifying AI risks, which governance Maps, Measures, and Manages.	Computational → Governance	USA
Entity-Based Regulation Framework [7]	Focused on transparency and regulating *"the large business entities developing the most powerful AI models and systems"* with emphasis on preemptive risk regulation.	Governance → Computational	USA
Three-Layered Framework [60]	Focused on risk, aims to fix market failures using a toolbox of regulatory tools from the three layers: market-invigorating strategies, value-directed rules, and procedural controls.	Computational → Governance	Global
UN System Survey of Institutional Models [52]	Focused on ethics and risk; highlights *"[national] capacity-building [that] can support AI development that is grounded in fairness, gender equality, reliability, safety, interpretability and accountability."*	Computational → Governance	UN/Global
Responsible AI Systems Roadmap [20]	Focused on risk, highly dependent on a committee of scientists to shape policy.	Computational → Governance	UN/Global
AI Ecological Education Policy Framework [10]	Focused on education; organized into 3 dimensions of educational support: pedagogical, ethical, and operational.	Computational → Governance	Hong Kong

Table 8. Existing governance frameworks recognize that there is both *computational capacity* and *governance capacity*.

Governance Capacity → Computational Capacity. The Hourglass Model of Organizational AI Governance [35] represents an ethics-first pipeline in which governance capacity flows toward computation. Ethical principles and organizational oversight mechanisms are translated into engineering practice. While there are feedback mechanisms for communication, the model largely presumes institutions can directly steer technical behavior. A similar directional logic appears in the Entity-Based Regulation Framework [7], which centers regulatory oversight of large corporate actors, confined to the level of enforcement and transparency rather than broader institutional design. Both frameworks view governance as an initiating force, with the Hourglass model supporting a limited feedback mechanism.

Computational Capacity → Governance Capacity. Other frameworks reverse this flow, positioning technical development as the driver of governance. The NIST AI Risk Management Framework [37] sequences accountability as mapping, measuring, and managing risk and emphasizes flexibility, voluntarism, and scalability across sectors:

> *"The Framework is intended to be voluntary, rights-preserving, non-sector-specific, and use-case agnostic, providing flexibility to organizations of all sizes and in all sectors and throughout society to implement the approaches in the Framework."*

The Three-Layered Framework [60] follows a similar pipeline but extends it globally, envisioning governance as a set of layered responses to computational markets – drawing from the toolbox categories of market-invigorating strategies, value-directed rules, and procedural controls. The UN System Survey of Institutional Models [52] emphasizes "capacity-building" but defines it primarily as the technical training of scientists and regulators.The Responsible AI Systems Roadmap [20] focuses on the role of scientists in shaping policy. These models recognize the importance of institutional development yet largely equate capacity with technical skill or regulatory compliance rather than with civic or educational infrastructure.

F.3 Institutional Capacity

Institutional Capacity extends beyond Governance Capacity to include the civic, educational, industrial, and governmental infrastructures that make computational work socially legitimate and accountable. It encompasses the systems and organizations that translate technical advancement into into legal systems, economic sectors, government, defense, and culture, and therefore requires a wide and interdisciplinary range of expertise:

> *"Questions about the impact of AI on American society and culture are fundamentally rooted in such humanities fields as ethics, law, history, philosophy, anthropology, sociology, media studies, and cultural studies."* –National Endowment for the Humanities [36]

While computational capacity has been extensively benchmarked, its rapid expansion has outpaced the institutional systems needed for accountability.

F.4 The Translation Workforce

The translation workforce is professionals who bridge technical and institutional domains. Algorithmic action is inherently situated – systems can have unintended outcomes when deployed in real social and organizational contexts [24, 49]. Effective governance therefore requires people and processes capable of translating between computational reasoning and institutional judgment.

The translation workforce goes beyond AI literacy (understanding how AI operates) or ethics (identifying harms). It encompasses the practical work of institutional integration: drafting regulations, designing oversight mechanisms, adjudicating disputes, negotiating standards, and explaining algorithmic decisions to diverse stakeholders, and/or exercising judgment about when AI implementation is not appropriate. Engineers, policymakers, educators, artists, cultural critics, and legal professionals all participate in this interpretive process.

G A Case Study on AI Detection Tools

Concerns about plagiarism and unauthorized use of generative AI tools are often treated as isolated classroom management issues. In reality, they reveal deeper failures of governance capacity – specifically, the lack of clear institutional policies, interpretive frameworks, and governance mechanisms capable of addressing AI use in equitable, accountable, and pedagogically meaningful ways. Universities have frequently responded to the rise of generative tools with restrictive policies or punitive enforcement regimes, but these approaches are often built on vague or inconsistent definitions of permissible use. The result is widespread uncertainty among both students and faculty [55], inconsistent application of standards, and in some cases, false accusations of academic misconduct.

While it is true that many faculty are concerned with academic dishonesty with the rise of generative AI, it is also the case that for at least some faculty, the concern goes beyond the (dis)honesty question and is rooted in a worry that students will fail to learn; as the AAUP Report puts it:

> *"The distinction between honesty and failure to learn is critical because it highlights one of the core goals of higher education: to develop a well-informed and thoughtful citizenry. This finding suggests that there is a need for higher education to refocus on the relational aspects of education and learning, as opposed to punitive measures..."* – Paris et al. [41]

Such outcomes erode trust, undermine institutional legitimacy, and highlight the urgent need for governance structures that are transparent, participatory, and aligned with civic values.

The stakes of false accusations extend far beyond individual classroom incidents. Let us consider Blackstone's Ratio [8]: "It is better that ten guilty persons escape than one innocent suffer." This principle remains instructive in the context of AI governance. Institutions that tolerate high rates of false positives in AI detection – for example, by punishing students based on unreliable systems – risk delegitimizing their authority and weakening the educational contract itself. Even a 1% false positive rate can result in thousands of wrongful accusations across large institutions, eroding trust in both faculty and administrative oversight, as illustrated in Figure 2. Protecting the innocent is foundational to sustaining the legitimacy of the institution as a whole.

Moreover, reliance on flawed detection technologies [59] introduces new forms of **procedural injustice**. Many detection systems disproportionately misclassify the work of multilingual students or those who write in nonstandard styles, compounding existing inequities [9, 27]. At the same time, the usefulness in certain contexts and ubiquity of generative AI virtually guarantee its continued presence in certain academic and professional contexts [6]. Attempts to suppress use via detection technologies are both impractical and counterproductive. Instead, educators should focus on equipping students with the critical skills needed to evaluate, contextualize, and responsibly integrate (or not integrate) AI-generated content into their work.

These challenges highlight that preventing misuse is not simply a question of enforcement but one of governance design. Effective institutional responses will require clear and consistently applied definitions of acceptable AI use, transparent policies with built-in safeguards such as appeal and review mechanisms, and curriculum that teaches students how to engage with generative AI ethically and productively. By shifting from a reactive, punitive stance to

Fig. 10. **False Positive Rate of 1%**: As an example, a false positive rate of only 1% [17, 51] indicates that 1 in every 100 cases will be flagged for cheating incorrectly. It falls on the instructor to evaluate use – faculty who are ill-equip to assess state-of-the-art AI systems. At scale in a university setting, for every 100,000 submissions, 1000 false accusations could occur, jeopardizing student careers and significantly de-legitimizing educational institutions.

a proactive governance model, universities can both uphold academic integrity and prepare students for meaningful participation in society increasingly pervaded by AI.

Finally, uneven access to generative tools and inconsistent usage policies can exacerbate existing inequities. Students with more resources, prior exposure, or permissive faculty gain advantages, while others face sanctions or are discouraged from engaging with technologies that are increasingly integral to professional and civic life. Limit and audit the use of AI-detection systems, which often produce false positives and disproportionately penalize multilingual and nontraditional writers. Any detection-based process must include transparent documentation, human review, and accessible appeals. Reliance on flawed detection software shows computational capacity outpacing institutional safeguards — a breakdown in procedural accountability.

H Public Website

To support transparency, inspection, and reuse of the ACAI-US79 audit, we provide a public, interactive website at acai-us79.org. The site is designed to make both the audit results and the audit instrument accessible beyond the paper, enabling readers to explore institutional AI governance capacity and to trace aggregate scores back to the underlying publicly available materials.

As shown in Figure 4, the primary interface is an interactive map-based visualization of the 79 U.S. universities included in the audit. Each institution is represented as a clickable marker, and a synchronized sidebar lists universities ranked by ACAI score. Users can filter institutions by research activity, institutional type, geographic region, and student size, and can select either a map marker or a list entry to reveal institution-specific details. For each university, the interface displays its ACAI score alongside direct links to the institutional policies, teaching resources, governance committees, and academic integrity materials reviewed during the audit. This design explicitly supports traceability, allowing users to inspect how scores are grounded in publicly legible governance artifacts rather than treating the index as an opaque ranking.

The website also includes a self-scoring interface for institutions not included in the ACAI-US79 dataset. This tool implements the same annotation schema and scoring criteria used in the audit, enabling users to assess their own institution using publicly available materials under comparable constraints. Self-scoring results are returned solely for informational and reflective purposes and are not incorporated into the released dataset.

The application is implemented in React with TypeScript, using `react-simple-maps` for geographic visualization and Material-UI for interface components. Institutional metadata—including location coordinates, research classification, institutional type, and associated governance resources—is stored in a structured JSON format derived from the original audit datasets. The interface supports interactive filtering, zooming, and state-level annotations to facilitate exploration and comparison across institutions.

Overall, the website is designed to emphasize interpretability and accountability. ACAI scores reflect the public legibility of institutional governance artifacts at the time of review and should not be interpreted as measures of internal practice or intent. By making both the audit results and the audit instrument publicly accessible, the site supports replication, critique, and institutional self-reflection, and positions ACAI as an auditable, contestable infrastructure rather than a static evaluation.

I Details on LLM-Driven Audit Study

Outputs failing schema or completeness checks were automatically retried until a valid response was produced.

Figure 11 shows the prompt used for the study.

```
Evaluating whether a university meets specific criteria based strictly on the provided
institutional links.

University: {UNIVERSITY}

Approved sources (you must not use any external sites, but you should explore sublinks
from these sites):
{LINK_LIST}

Evaluation criteria:
{CRITERIA_LIST}

Scoring scale (select exactly one per criterion):
A. Present/Yes — A clear statement directly addressing the item is found on an institutional
page within 5 minutes.
B. Partial/Implicit/Somewhat — The item is mentioned or implied, but key details are missing.
C. Absent/No — You reasonably searched the allowed sources and did not find relevant content.
D. Unclear or Took Longer Than 5 Minutes — Navigation difficulty, vague language, or time limits
prevented a confident decision.
E. Conflicting Information — Different institutional sources provide contradictory guidance for
the same item.

Output requirements:
- Return VALID JSON ONLY.
- No markdown, no commentary.
- Return an ARRAY with one object per criterion.
- Every criterion MUST appear exactly once.

Schema:
[
    {
        "criterion": "A1",
        "score": "A|B|C|D|E",
        "urls": ["https://...", "https://..."]
    }
]
```

Fig. 11. **Prompt used in the LLM robustness study.** The figure shows the full evaluation prompt provided to the language model, including source restrictions, scoring criteria, and the required JSON output schema, mirroring the instructions given to human annotators in the ACAI audit.