Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The University AI Sovereignty Index: A Multidimensional Method for Assessing Institutional Digital Autonomy

Mustafa Kayyali

Kayyali@heranking.com

Maaref University of Applied Sciences

Research Article

Keywords: artificial intelligence, digital sovereignty, higher education, composite index, institutional autonomy, AI governance, quality assurance

Posted Date: November 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11451660/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

The University AI Sovereignty Index: A Multidimensional Method for Assessing Institutional Digital Autonomy

Highlights

- UASI measures practical AI autonomy across eight institutional dimensions.
- Evidence-adjusted indicators distinguish documented capability from aspiration.
- Geometric aggregation prevents a strong domain from concealing a critical weakness.
- OCAET makes educational legitimacy part of sovereignty rather than an afterthought.
- The index is diagnostic and requires empirical validation before comparative use.

Abstract

Universities are adopting artificial intelligence through cloud platforms, enterprise assistants, learning systems, research tools, and administrative analytics, yet most institutions lack a method for judging whether these arrangements preserve meaningful digital autonomy. This methodological paper develops the University AI Sovereignty Index (UASI), a multidimensional assessment of an institution's practical capacity to choose, govern, scrutinise, and, when necessary, exit AI dependencies. The index was constructed through an integrative synthesis of digital sovereignty, platform governance, responsible AI, learning analytics ethics, composite-indicator methodology, and higher-education policy research. UASI contains eight dimensions and 40 evidence-rated indicators: strategic authority, infrastructural control, data stewardship, model assurance, pedagogical autonomy, human capability, ecosystem negotiability, and collective rights. Scores use a weighted geometric aggregation that limits compensation between dimensions, supplemented by critical safeguards for high-consequence systems. The OCAET test - orient knowledge, activate cognition, preserve agency, support evaluation, and enable transfer - anchors the pedagogical dimension. Worked hypothetical cases show why an institution with strong infrastructure but weak contestability should not receive the same judgment as a balanced institution with a similar arithmetic mean. The paper provides a scoring rubric, evidence protocol, sensitivity procedure, reporting dashboard, and staged validation programme. UASI is presented as a falsifiable methodological proposition rather than a validated ranking. Its purpose is diagnostic: to make hidden dependencies discussable, reveal capability gaps, and support better institutional choices without confusing sovereignty with technological isolation.

Keywords: artificial intelligence; digital sovereignty; higher education; composite index; institutional autonomy; AI governance; quality assurance

Plain Language Summary

As universities use AI to inform critical decisions, the need to understand and evaluate its adoption (UASI) becomes paramount. Just like a university might embrace AI in many advanced ways and simultaneously enter into a deal for a model with little control or insight about the data that was used, UASI provides a structured and systematic way to examine this problem. UASI assesses a university's understanding and evaluation of AI adoption in eight areas. It encourages leadership authority over infrastructure and continuing education of the staff as well as rewarding balanced capability across all 8 areas. Hypothetical examples of performance are provided for explanation; real institutions should validate this instrument before relying on it for comparison or public claims.

1 Introduction

Independent digital sovereignty is a rare achievement for universities. While the academic community at an institution can aspire toward autonomy in their intellectual and operational activities, the systems and services that support universities are capable of captive dependency. The potential for control of research and teaching is hardly limited to dystopian fables. An academic staff member determines and becomes accustomed to a data collection schema, then a learning platform introduces some integration opportunities that entrench and widen the schema, and finally, during a routine process of renewing a software licence, an AI assistant is added. Over the years, countless such arrangements, made apparently without great consequence, eventually erode a university's autonomy over its data, model, and learning architecture. Few committees are ever required to read the terms of service of downstream providers or understand the consequences of widening the data schema. Yet they ultimately bind an institution into relationships that are at least difficult and often impossible to escape. Digital sovereignty to some might not be raising a flag in a virtual territory but rather having the capacity to exercise control over what is consequential.

The concept of digital sovereignty has recently inspired discussions about control on a national scale (Couture & Toupin, 2019; Pohle & Thiel, 2020) and at the level of the individual (Floridi, 2020). As applied to the environment of universities, digital sovereignty would include data governance, control over infrastructure, and capacity for self-determined development. However, while digital sovereignty might be an appropriate framing for some aspects of university technology usage, it is critique-prone because it carries all those meanings at once. A university could build a 'sovereign cloud' and still be epistemically dependent on predictions made by proprietary models. Likewise, universities adopting open-source software and hosting it on local servers do not become sovereign by mere ownership unless they own sufficient capacity to operate and secure the system. Similarly, procuring services in compliance with data protection legislation does not automatically bestow sovereignty if there is no possibility of refusing prepackaged technology.

Many recent frameworks for governing AI systems offer useful insights for developing a clear vision of digital sovereignty (Autio et al., 2024; Miao & Holmes, 2023; National Institute of Standards and Technology, 2023; UNESCO, 2021). The NIST framework for AI risk management, for example, clearly differentiates between activities for governing, mapping, measuring, and managing risk. The unconditional governance of generative AI is an emerging but currently unrealistic ideal, and a framework for managing risk would be an appropriate goal for most producers and users of AI technologies. Similarly, UNESCO suggests a carefully crafted measurement architecture, but most framework implementations so far resemble university-policy studies that propose a blanket prohibition on all AI applications or alternatively welcome their conditional use, leaving substantial discretion to individual faculty members as to what conditions apply (An et al., 2025; Jin et al., 2025; Li et al., 2026; McDonald et al., 2025). Without a consistent and robust structure to guide implementation, even those frameworks leave users without the information necessary to assess their progress.

While composite indices are aggregated numbers that depend on all of the choices made along the way, their weightings can obscure the designer's political judgment (Nardo et al., 2008). Not all dimensions are created equal, and it is important to consider carefully the information that is lost through aggregation. The work of constructing a composite index also presumes that weakness in one dimension can be compensated for by strength in another, and many choices made in constructing the index and interpreting its results can have significant and unacknowledged effects. Further, evidence that is missing can affect a composite index in at least three distinct ways: it can be incorrectly assumed to be absent (resulting in unwarranted credit to the university), it can be correctly assumed to be absent (resulting in a penalty), or it can be imputed (resulting in a falsification that allows for a stronger index). Finally, composite indices can be published in a variety of formats, including league tables, and they can be used as reputational weapons to punish or reward whoever happens to be in the crosshairs (Saltelli, 2007).

We separate the score from the profile to avoid the above traps, and we commit ourselves to the integrity of the evidence and the measure itself.

This paper presents the University AI Sovereignty Index (UASI) as a practical structure for evaluating university sovereignty in the age of AI. The index score, which is inseparable from a profile, is intended not as an evidence grade in its own right but rather a measure of

critical safeguards, such as the existence of a policy on the risks of AI use and a commitment to monitor, document, and address these risks. The safeguards indicate areas where the ranking is merely an expression of a narrative account of dependencies rather than a serious engagement with the problem of AI governance for universities.

This paper has five purposes. First, we define university AI sovereignty as a measurable capability without reducing it to a binary autarky. This allows us to develop a multidimensional construct from which an indicator system might be defined. Second, we describe the aggregation method used in the index, which aims to avoid compensation at least in theory, and we discuss two possible measures of evidentiary confidence. Third, we demonstrate interpretation of the index score using a hypothetical case of a university with a predetermined profile. Fourth, we briefly describe an empirical validation programme to revise or reject the proposed architecture. Finally, we present a new framework for assessing AI systems designed for university use — Objective Characteristic Assessment for Educational Technology (OCAET) — as an educationally oriented test for AI systems. The test considers that sovereign action requires AI systems that orient knowledge for users rather than end up restricting it, activate cognition and preserve the user's agency instead of undermining it, support the user's evaluation of their work, and enable transfer to other activities as desired.

2 Materials and Methods

2.1 Construct definition and design principles

University AI sovereignty is defined here as the legitimate and practical capacity of a higher-education institution, exercised with affected communities, to choose the purposes, infrastructures, data relations, model conditions, and termination pathways through which AI enters academic and administrative life. Three words carry most of the burden. Legitimate means that institutional control is bounded by rights, shared governance, law, and public purpose. Practical distinguishes formal contractual rights from capabilities that can be used. Capacity points to an organisational accomplishment: authority, expertise, information, infrastructure, and bargaining power must be assembled before autonomy can be exercised.

Four design principles follow. The first is multidimensionality: sovereignty is legal, infrastructural, operational, epistemic, pedagogical, and relational. The second is non-compensation: excellent cybersecurity cannot erase the absence of appeal, just as robust teaching guidance cannot compensate for uncontrolled data extraction. The third is evidence sensitivity: a policy promise and an audited practice are not equivalent. The fourth is diagnostic restraint: UASI should direct inquiry and improvement before it is used for external comparison.

2.2 Evidence synthesis and item generation

Item generation followed an integrative, theory-synthesis procedure (Jaakkola, 2020; Snyder, 2019) appropriate for this phase of developing an emerging construct. It synthesised concepts and practices across six literatures: digital and data sovereignty; platform and infrastructure studies; responsible-AI governance and auditing; higher-education AI policy; learning-analytics ethics; and composite-indicator construction. Candidate items were retained for inclusion if they described an institutional decision or capability, were actionable and supported by inspectable evidence, and provided information that could not be more precisely captured elsewhere.

This yielded a pool of 67 candidate statements, from which those that were conceptually overlapping were removed, compound items were separated, and those that expressed purely aspirational language were translated into indicators of observable practices. This reduction generated 40 indicators across eight dimensions. The aim of this exercise was not to demonstrate content validity, but to propose a set of indicators. A formal Delphi and content-validity stage, yielding expert consensus, is planned for a later stage of this work, as authors should not simulate agreement after the fact. Likewise, weights assigned to the indicators are intended as a defensible starting theory, rather than an order necessarily discovered.

2.3 Scoring scale and evidence grades

Each indicator receives a practice score from 0 to 4. Zero indicates absence or a contradiction of the practice; 1 indicates an ad hoc or person-dependent response; 2 denotes a documented but partial process; 3 represents an implemented, monitored process with named ownership; and 4 requires an embedded, independently reviewable practice that generates evidence for revision. Scorers also assign an evidence grade: 0.50 for self-description without supporting artefacts, 0.75 for internally corroborated documentation, and 1.00 for externally verifiable or independently audited evidence. The evidence grade does not declare truth. It records how much confidence the assessment can reasonably bear.

Indicator values are calculated as the practice score multiplied by the evidence grade. A dimension score is the mean adjusted item value divided by four and multiplied by 100. Institutions must report missing items rather than treating them as zero. When more than one of five indicators in a dimension lacks usable evidence, the dimension is marked insufficient and no overall score is produced. This rule is intentionally inconvenient. A sovereignty assessment that rewards opacity would be conceptually perverse.

Table 1. Proposed UASI dimensions and starting weights

Dimension	Weight percent	Construct emphasis
Strategic authority	15	Mandate, named accountability, decision rights, risk appetite, board or senate oversight
Infrastructural control	15	Architecture visibility, security, configurability, resilience, replaceability
Data stewardship	15	Purpose limitation, lineage, minimisation, access, retention and deletion
Model assurance	15	Model documentation, local evaluation, subgroup analysis, monitoring, change control
Pedagogical autonomy	15	OCAET alignment, academic discretion, assessment validity, alternatives, learning evidence
Human capability	10	Leadership literacy, technical expertise, educator development, student literacy, workload
Ecosystem negotiability	10	Procurement leverage, portability, interoperability, concentration risk, tested exit
Collective rights	5	Participation, accessibility, contestability, community effects, public transparency

Note. Weights are theoretical starting values and must be tested through expert elicitation and sensitivity analysis.

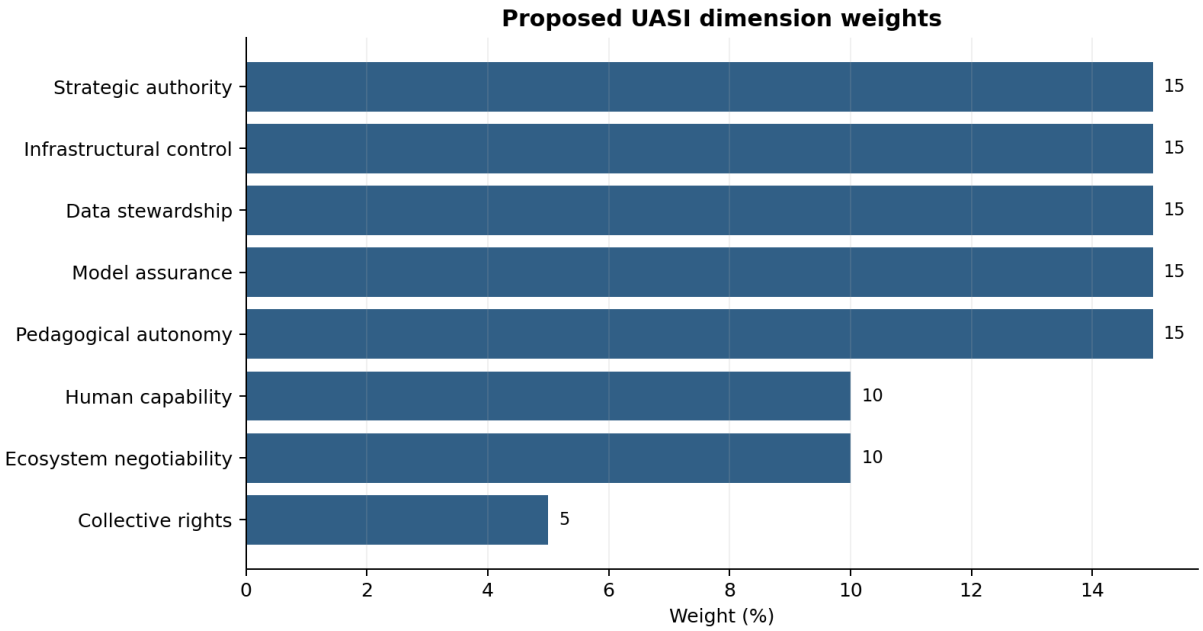


Figure 1. Proposed weighting architecture for the University AI Sovereignty Index

Note. The distribution prevents governance from becoming a purely technical or purely pedagogical measure.

2.4 Aggregation and critical safeguards

Let Dj denote a dimension score between 0 and 100 and wj its weight, with the weights summing to one. The base UASI is the weighted geometric mean: UASIbase = product over j of (Dj/100) raised to wj, multiplied by 100. Geometric aggregation rewards balance and makes low dimensions consequential. Unlike a simple arithmetic mean, it does not permit a spectacular infrastructure score to wash away weak collective rights or pedagogical autonomy. A score of zero in any dimension produces an overall score of zero because complete absence is treated as structural failure rather than a small deficit.

The number is then read through five critical safeguards: a named accountable owner; lawful and purpose-specific data authority; documented human review and appeal; tested incident suspension; and local validation for any high-consequence model. Failure of a safeguard does not alter the diagnostic profile, but it caps the institution at the 'constrained' category for high-consequence deployment. This dual system separates broad capability from permission to proceed. Readiness for an AI writing assistant, after all, is not readiness for automated admissions or progression decisions.

Table 2. Indicator scoring rubric

Score	State	Operational interpretation
0	Absent or contradictory	No evidence of the capability; practice may undermine it.
1	Ad hoc	Individuals improvise; continuity depends on personal effort.
2	Documented	A process exists but coverage, implementation, or monitoring is incomplete.
3	Implemented	Named ownership, routine operation, monitoring, and corrective action are visible.
4	Embedded and reviewable	Independent scrutiny, stakeholder evidence, learning, and revision are institutionalised.

3 Results

3.1 The eight-dimensional sovereignty architecture

3.1.1 Strategic authority

To hold strategic authority over the use of AI, a university must be able to answer the question: who governs AI, and to what extent, before a dispute occurs? A university AI committee with no delegated authority cannot stop a deployment, but convenes discussion around AI use. Central control that doesn't involve academics or students risks sovereignty being turned into a kind of managerial enclosure. Scoring highly under this criteria means that decision rights are made explicit, as well as having oversight from governing bodies and academia; a stated appetite for risk and a route for exceptions and escalation.

3.1.2 Infrastructural control

In discussing infrastructural control I am concerned with visibility and manoeuvrability with AI, not owning the server on which it runs. Knowing the dependencies, having the ability to configure them, stopping the application, maintaining continuity when it goes down, the ability to replace components; the ability to do those things gives us control. If we aspire to replace the AI infrastructures so the educational environment gives us the one we want, we need the ability to do all of these things (Plantin et al., 2018; van Dijck et al., 2018).

3.1.3 Data stewardship

To score highly under this criteria, universities would govern the data as contextual and relational material, with respect to the lineage of data, purpose limitation, minimisation, access control, retention, the ability to verify deletion, restrictions on model training, and so on. Individual privacy matters, but so too do the collective effects of data use. The CARE principles say that the governance of data should include considerations of authority, responsibility, ethics and collective benefit (Carroll et al., 2020).

Model assurance: The epistemic layer is about what's possible and what's sought. It includes documentation, locally relevant testing, subgroup performance, uncertainty, monitoring, material change control. A model card from a vendor gives some information, but reveals nothing about whether our engagement score was determined through learning in a particular programme, or how well a multilingual assistant is likely to serve the university's students. Our assurance needs to be constructed and tested in context.

Pedagogical autonomy is where UASI departs most clearly from the generic digital-sovereignty indices. While OCAET asks about AI use orientation, knowledge activation, cognitive preservation, agency preservation, and evaluation and transfer enablement, it makes no direct connections to how those assurances would play out in practice. However, our pedagogy informs how we will use AI, and how we will ensure each of those purposes is met as well as possible. Owning our own models, but designing an education system in which students design questions and then passively await the answers, is an example of technical control without educational sovereignty. Human capability assesses how well people are able to exercise the rights and accept the responsibilities assigned to them. This includes leadership literacy, technical and legal expertise, developer and educator development, student AI literacy, opportunities for protected challenge, and workload. All governance structures collapse if responsibility is delegated while the time, tools, and support to carry it out remain centralised elsewhere. Policy research is increasingly identifying this gap between discretion and capacity (e.g. Azevedo et al., 2025; McDonald et al., 2025).

3.2 The OCAET sovereignty test

Table 3. OCAET as the educational legitimacy component of UASI

Purpose	Evidence of sovereignty	Failure pattern
Orient knowledge	Provenance, limitations, uncertainty, and model role are intelligible.	Fluent output is treated as settled knowledge.
Activate cognition	AI interaction prompts analysis, comparison, diagnosis, or revision.	The system performs the intellectual operation being assessed.
Preserve agency	Humans retain meaningful choice, authorship, refusal, and review.	Use is nominally optional but practically compulsory.
Support evaluation	Claims and outputs can be checked against disciplinary criteria.	Accuracy or benefit is inferred from vendor benchmarks alone.
Enable transfer	Learning and institutional capability survive a change of tool or context.	Knowledge and workflow remain locked inside the system.

OCAET is scored through artefacts rather than declarations: assessment briefs, moderation records, student-facing guidance, accessible alternatives, examples of verification, delayed or independent performance evidence, and documented pedagogical review. The purpose is not to demand that every AI use maximise all five functions. A translation aid may chiefly preserve agency; a critique exercise may chiefly activate cognition and support evaluation. Across a programme, however, persistent weakness indicates that technology is substituting for the capacities education claims to develop.

3.3 Worked hypothetical application

Three fictional institutions illustrate interpretation. Meridian University has substantial infrastructure, data controls, and technical expertise, but weak participation and limited pedagogical review. Commons University has moderate technical resources and unusually strong collective governance, teaching design, and exit planning. Frontier College is innovative at course level but depends on free consumer systems and person-specific expertise. The figures are not observations and should never be presented as evidence about real universities.

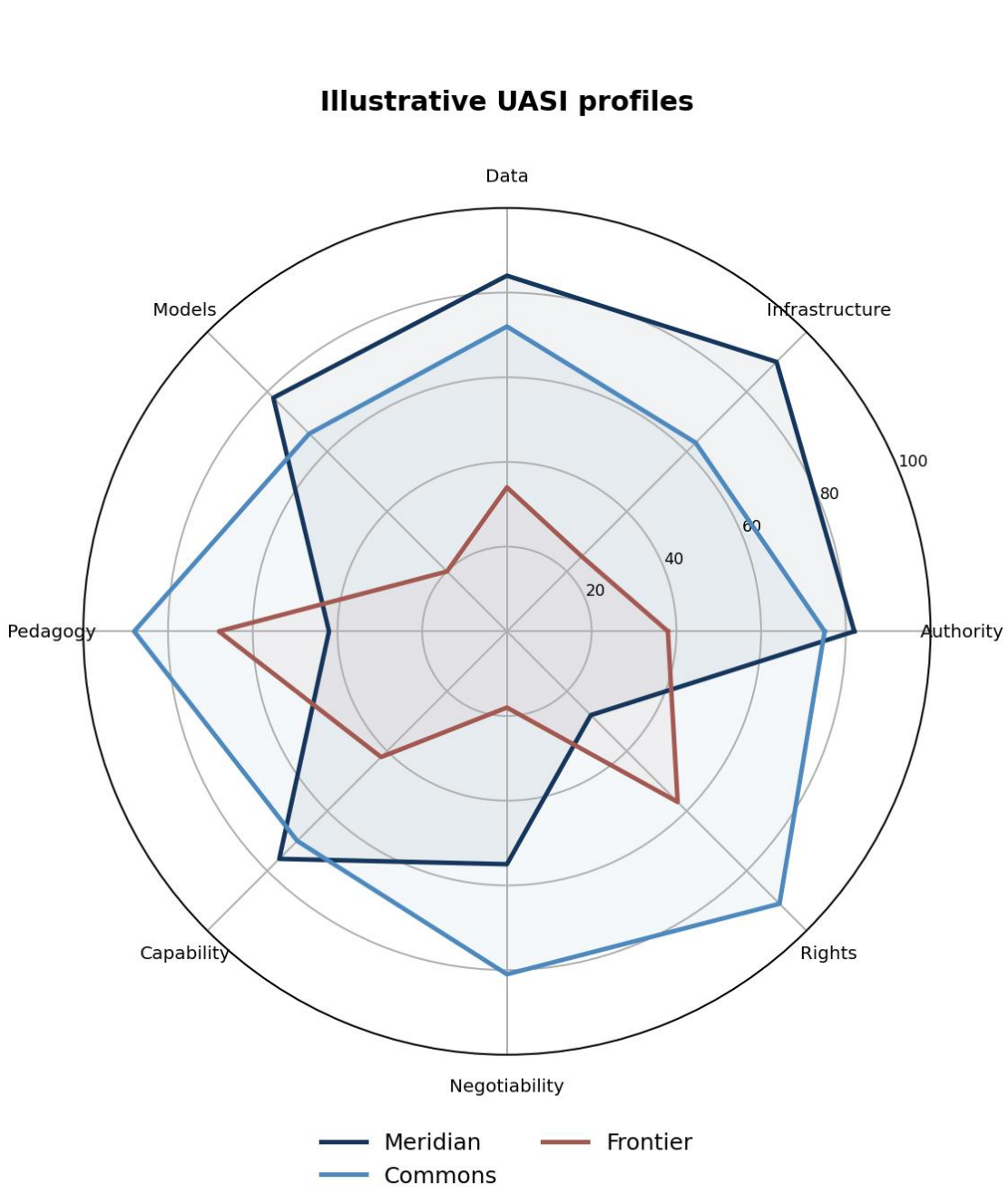


Figure 2. Hypothetical institutional profiles across UASI dimensions

Note. Values are invented solely to demonstrate interpretation; they are not empirical results.

Table 4. Interpretation of the hypothetical cases

Institution	Illustrative UASI	Interpretive judgment
Meridian University	66.0	Strong technical capability but constrained by weak rights and pedagogical autonomy.
Commons University	74.7	Balanced profile; improvement priority is infrastructural and model assurance.
Frontier College	32.4	Creative local use rests on fragile infrastructure, data, assurance, and exit capacity.

Note. Scores were calculated from the displayed hypothetical profiles using the proposed geometric method; rounding is illustrative.

The cases expose the weakness of arithmetic comfort. Meridian's average looks respectable, yet the institution's rights score means that affected people may have little influence or redress. Commons scores higher not because it dominates every dimension but because its capabilities cohere. Frontier should not be instructed simply to

'adopt more AI'. Its immediate task is to make dependencies visible, establish data rules and accountable ownership, and secure supported alternatives before scaling consequential uses.

4 Discussion

4.1 What the index contributes

Because UASI reframes sovereignty not as symbolic independence, but as governed dependency, it matters most — particularly for universities, as interdependence is constitutive of scholarship — whether that dependence is mediated through cloud services, international research networks, open-source communities, public infrastructures or commercial platforms. It must be visible, negotiable, reviewable and compatible with educational purpose.

The index also makes explicit the hand-offs that must be respected in responsible management of automated systems: between disciplinary domains that might otherwise be governed apart (procurement measures price as well as security; committees assess teaching, data-protection offices lawful processing, IT integration; students balance whether use is fair or compulsory) and between professions whose expertise cannot just be rolled together (UASI does not imagine it difficult to dissolve problems of risk analysis into a single office).

The non-compensatory logic of the index also expresses a university's commitment to governance of automated decisions that does not compromise on any of its threshold conditions. A university with any appeal or suspension on any of the individual dimensions of the index will not be ready to take the risk of making high-consequence automated decisions — even if its computing environment is otherwise fairly impressive.

4.2 Risks of measurement

All such indices risk becoming the institution that they wish to observe, especially once there is a reputation or funding attached to a score. Here, the documentation, or actuality, of a particular capability may be an advantage; hiding incidents may be a disadvantage; and the measurement of controls may lead to a change of measurement. All may apply even more strongly when public ranked comparison is introduced.

Given these serious risks, UASI should be deployed, first and foremost, in a confidential developmental context. In that context, notes generated from evidence should be subjected to intensive review by stakeholders. Only then should the complex process of external comparison be contemplated, and only after measurement invariance, audit reliability, and contextual fairness have been demonstrated.

The politics of weighting is best illustrated through the proposed architecture, with 15 percent to each of the five core domains, 10 percent to capability and negotiability, and 5 percent to collective rights. A different community may allocate greater weight to rights, however, or even to infrastructure, if able to work in consort. As noted, sensitivity analysis, reporting rankings and classifications emerging from equal weights, expert weights, and other plausible perturbations illuminate whether uncertainty is inherent in the conclusion or easily reversible.

The index can be applied even in settings where resource privilege does not exist. For example, a small university may not have dedicated staff to develop AI systems in house, but may choose to have strong autonomy in the way it uses public infrastructure as long as it is able to gain assurance about its consortial use. The indicators are written to reflect effective capability rather than institutional ownership, but would also work for systems provided collaboratively. Evidence collection may be more burdensome and this will need to be adapted with, for example, a shorter form for reports, sector support for peer-review of evidence, and piloting of the exchange.

4.3 Validation programme

Table 5. Proposed empirical validation programme

Phase	Procedure	Decision output
1 Construct clarification	Cognitive interviews with university leaders, educators, students, privacy officers, technologists, and procurement specialists.	Revised definitions and item wording.
2 Content validity	Two-round international Delphi; report item-level CVI, Aiken's V, disagreement, and regional representation.	Retained, revised, and rejected items.
3 Feasibility pilot	Dual-rater assessment in 30-50 institutions varying by region, mission, size, and resources.	Burden, missingness, evidence availability, and inter-rater reliability.
4 Structural validation	Larger sample; test dimensionality, internal consistency, discriminant validity, and measurement invariance.	Empirically supported or revised architecture.
5 Criterion and predictive validity	Relate profiles to audit findings, incidents, exit performance, stakeholder trust, and educational outcomes.	Evidence that the instrument predicts consequences rather than paperwork.
6 Longitudinal calibration	Repeat assessments after policy, vendor, or system changes.	Responsiveness, stability, and minimal meaningful change.

Content validity should use a deliberately plural panel. An index designed only by technology and legal experts would miss pedagogy; one designed only by educators might under-specify infrastructure and contracts. Item-level agreement should be reported, not hidden inside a global coefficient (Aiken, 1985; Lawshe, 1975). Later psychometric work can examine internal structure, but high alpha is not the goal. Dimensions such as rights and infrastructure should relate without collapsing into one trait (Cronbach, 1951; DeVellis & Thorpe, 2021).

Criterion validation presents a harder problem because there is no gold standard for sovereignty. A portfolio of external criteria is preferable: verified incidents, time to suspend a system, success of data export, contract changes achieved, stakeholder understanding, appeal outcomes, model performance drift, and evidence of learning. Convergent evidence across these outcomes would be more persuasive than correlation with another maturity score.

4.4 Implications for policy and quality assurance

By focusing more on the profile of the institution rather than one set of technology, quality-assurance bodies will be better able to ask sharper questions. For example:

- Does the institution know which systems are used to shape decisions for students?
- What does the institution know about locally evaluated high-consequence models?
- Are course-level permissions coherent with institutional privacy rules?
- How has AI-supported learning been shown to transfer beyond the supported task?

Visibility can be a score, or it can be a step towards something more. For institutional leaders who have worked through the sequence we recommend, scoring theatre is not the point. Assuring a data and model register, orienting a decision-right map, and marshalling the OCAET to review high-impact educational uses can all be rehearsed and exited on a small scale—and reveal much—before a public score is released. We anticipate that as evidence practices stabilise, UASI will serve both as a guide to annual improvement and a spur to strategic procurement. Used in this way, the index can be invaluable. Better to start low and have a conversation than end high and close it.

5 Conclusion

Approached carefully, the University AI Sovereignty Index offers a systematic method for examining whether universities most fundamental quality is being compromised by our expanding dependencies on AI. The index sets out eight dimensions of strategic authority, assuring the infrastructure, securing the data, developing models,

enhancing pedagogy, growing human capability, understanding our ecosystem's negotiability and preserving collective rights. Graded against these dimensions, practices can be distinguished from promises and complicit maths avoided through geometric aggregation. Safeguards placed on critical dimensions ensure that high-score is not delivered alongside high-consequence deployment. Keeping the OCAET in review lock by steering the assessment towards education holds all this together. Yet because the UASI is not validated as a ranking with all-dimensional review and empirical work to reject or confirm the weights, it would be methodologically imprudent to claim it can accomplish any of these functions with real certainty. We hope that others will take up the challenge to create better measures. These types of sovereignty are not demonstrated by owning everything. They are demonstrated by understanding and governing through dependency, maintaining the role for human judgment, being open to being negotiated with, and retaining the ability to leave without significant sacrifice..

Appendix A Proposed UASI indicator specification

Table A1. Forty proposed indicators and evidence families

Dimension	Item	Indicator	Illustrative evidence
Strategic authority	1	Approved institutional AI mandate	Policy; artefact; log; audit; stakeholder evidence
	2	Named accountable executive and governance body	Policy; artefact; log; audit; stakeholder evidence
	3	Published decision-right and escalation map	Policy; artefact; log; audit; stakeholder evidence
	4	Risk appetite differentiated by consequence	Policy; artefact; log; audit; stakeholder evidence
	5	Governing-board and academic-governance review	Policy; artefact; log; audit; stakeholder evidence
Infrastructural control	1	Maintained AI system and dependency register	Policy; artefact; log; audit; stakeholder evidence
	2	Configurable identity, access, and logging	Policy; artefact; log; audit; stakeholder evidence
	3	Security and resilience testing	Policy; artefact; log; audit; stakeholder evidence
	4	Modular interfaces and replaceability	Policy; artefact; log; audit; stakeholder evidence
	5	Documented continuity arrangements	Policy; artefact; log; audit; stakeholder evidence
Data stewardship	1	Purpose and lawful-basis register	Policy; artefact; log; audit; stakeholder evidence
	2	Data-lineage and subprocessor visibility	Policy; artefact; log; audit; stakeholder evidence
	3	Minimisation and role-based access	Policy; artefact; log; audit; stakeholder evidence
	4	Retention and verified deletion	Policy; artefact; log; audit; stakeholder evidence
	5	Restrictions on secondary training and reuse	Policy; artefact; log; audit; stakeholder evidence
Model assurance	1	Model and capability documentation	Policy; artefact; log; audit; stakeholder evidence
	2	Local task and subgroup evaluation	Policy; artefact; log; audit; stakeholder evidence
	3	Uncertainty and limitation communication	Policy; artefact; log; audit; stakeholder evidence
	4	Drift, incident, and outcome monitoring	Policy; artefact; log; audit; stakeholder evidence
	5	Material-change notification and reapproval	Policy; artefact; log; audit; stakeholder evidence
Pedagogical autonomy	1	Orient knowledge	Policy; artefact; log; audit; stakeholder evidence
	2	Activate cognition	Policy; artefact; log; audit; stakeholder evidence

Dimension	Item	Indicator	Illustrative evidence
	3	Preserve agency	Policy; artefact; log; audit; stakeholder evidence
	4	Support evaluation	Policy; artefact; log; audit; stakeholder evidence
	5	Enable transfer	Policy; artefact; log; audit; stakeholder evidence
Human capability	1	Leadership and board AI literacy	Policy; artefact; log; audit; stakeholder evidence
	2	Technical, legal, and assurance expertise	Policy; artefact; log; audit; stakeholder evidence
	3	Disciplinary educator development	Policy; artefact; log; audit; stakeholder evidence
	4	Student AI and verification literacy	Policy; artefact; log; audit; stakeholder evidence
	5	Workload and protected challenge mechanisms	Policy; artefact; log; audit; stakeholder evidence
Ecosystem negotiability	1	Contractual audit and change rights	Policy; artefact; log; audit; stakeholder evidence
	2	Usable data and workflow portability	Policy; artefact; log; audit; stakeholder evidence
	3	Interoperable, documented interfaces	Policy; artefact; log; audit; stakeholder evidence
	4	Provider-concentration assessment	Policy; artefact; log; audit; stakeholder evidence
	5	Tested exit and migration plan	Policy; artefact; log; audit; stakeholder evidence
Collective rights	1	Early stakeholder participation	Policy; artefact; log; audit; stakeholder evidence
	2	Accessible alternatives and accommodations	Policy; artefact; log; audit; stakeholder evidence
	3	Human review and appeal	Policy; artefact; log; audit; stakeholder evidence
	4	Collective and community impact assessment	Policy; artefact; log; audit; stakeholder evidence
	5	Public transparency and version history	Policy; artefact; log; audit; stakeholder evidence

Note. The full rubric should be refined through cognitive testing and Delphi review before field use.

Declarations

Author contributions: The author was responsible for conceptualisation, methodology, analysis, visualisation, and preparation of the manuscript.

Funding: No external funding was received for this study.

Conflict of interest: The author declares no conflict of interest.

Ethics approval: Ethical approval was not required because the study did not involve human participants, identifiable personal data, or clinical material.

Data availability: No new empirical dataset was generated. The conceptual materials, scoring specifications, and evidence sources used in the study are reported in the manuscript and its tables.

Generative artificial intelligence disclosure: Generative AI tools were used extensively during development of this draft to support synthesis, drafting, language development, and document preparation.

References

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. Educational and Psychological Measurement, 45(1), 131-142. https://doi.org/10.1177/0013164485451012

An, Y., Yu, J. H., & James, S. (2025). Investigating the higher education institutions' guidelines and policies regarding the use of generative AI in teaching, learning, research, and administration. International Journal of Educational Technology in Higher Education, 22, Article 10. https://doi.org/10.1186/s41239-025-00507-3

Autio, C., Schwartz, R., Dunietz, J., Jain, S., Stanley, M., Tabassi, E., Hall, P., & Roberts, K. (2024). Artificial intelligence risk management framework: Generative artificial intelligence profile (NIST AI 600-1). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.600-1

Azevedo, L., Robles, P., Best, E., & Mallinson, D. J. (2025). Institutional policies on artificial intelligence in higher education: Frameworks and best practices for faculty. New Directions for Adult and Continuing Education, 2025(188), 70-78. https://doi.org/10.1002/ace.70013

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 610-623. https://doi.org/10.1145/3442188.3445922

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. Frontiers in Public Health, 6, Article 149. https://doi.org/10.3389/fpubh.2018.00149

Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. Assessment & Evaluation in Higher Education, 31(4), 399-413. https://doi.org/10.1080/02602930600679050

Carroll, S. R., Garba, I., Figueroa-Rodriguez, O. L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., Walker, J. D., Anderson, J., & Hudson, M. (2020). The CARE principles for Indigenous data governance. Data Science Journal, 19, Article 43. https://doi.org/10.5334/dsj-2020-043

Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. International Journal of Educational Technology in Higher Education, 20, Article 38. https://doi.org/10.1186/s41239-023-00408-3

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. International Journal of Educational Technology in Higher Education, 20, Article 43. https://doi.org/10.1186/s41239-023-00411-8

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innovations in Education and Teaching International, 61(2), 228-239. https://doi.org/10.1080/14703297.2023.2190148

Couture, S., & Toupin, S. (2019). What does the notion of sovereignty mean when referring to the digital? New Media & Society, 21(10), 2305-2322. https://doi.org/10.1177/1461444819865984

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16, 297-334. https://doi.org/10.1007/BF02310555

Dawson, P. (2021). Defending assessment security in a digital world: Preventing e-cheating and supporting academic integrity in higher education. Routledge.

DeVellis, R. F., & Thorpe, C. T. (2021). Scale development: Theory and applications (5th ed.). SAGE.

Drachsler, H., & Greller, W. (2016). Privacy and analytics: It's a DELICATE issue. Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, 89-98. https://doi.org/10.1145/2883851.2883893

European Commission. (2022). Ethical guidelines on the use of artificial intelligence and data in teaching and learning for educators. Publications Office of the European Union. https://doi.org/10.2766/153756

European Parliament & Council of the European Union. (2016). Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data. Official Journal of the European Union, L 119, 1-88.

European Parliament & Council of the European Union. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence. Official Journal of the European Union, L 2024/1689.

Floridi, L. (2020). The fight for digital sovereignty: What it is, and why it matters, especially for the EU. Philosophy & Technology, 33, 369-378. https://doi.org/10.1007/s13347-020-00423-6

Hummel, P., Braun, M., Tretter, M., & Dabrock, P. (2021). Data sovereignty: A review. Big Data & Society, 8(1). https://doi.org/10.1177/2053951720982012

Ifenthaler, D., & Schumacher, C. (2016). Student perceptions of privacy principles for learning analytics. Educational Technology Research and Development, 64, 923-938. https://doi.org/10.1007/s11423-016-9477-y

Jaakkola, E. (2020). Designing conceptual articles: Four approaches. AMS Review, 10, 18-26. https://doi.org/10.1007/s13162-020-00161-0

Jin, Y., Yan, L., Echeverria, V., Gasevic, D., & Martinez-Maldonado, R. (2025). Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. Computers & Education: Artificial Intelligence, 8, Article 100348. https://doi.org/10.1016/j.caeai.2024.100348

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1, 389-399. https://doi.org/10.1038/s42256-019-0088-2

Jöhnk, J., Weißert, M., & Wyrtki, K. (2021). Ready or not, AI comes: An interview study of organizational AI readiness factors. Business & Information Systems Engineering, 63, 5-20. https://doi.org/10.1007/s12599-020-00676-7

Kasneci, E., Sessler, K., Kuchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Gunnemann, S., Hullermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

Lawshe, C. H. (1975). A quantitative approach to content validity. Personnel Psychology, 28(4), 563-575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x

Li, M., Xie, Q., Enkhtur, A., Meng, S., Chen, L., Yamamoto, B. A., Cheng, F., & Murakami, M. (2026). A framework for developing university policies on generative AI governance: A cross-national comparative study. Studies in Higher Education. Advance online publication. https://doi.org/10.1080/03075079.2026.2696496

Lodge, J. M., Howard, S., Bearman, M., Dawson, P., & Associates. (2023). Assessment reform for the age of artificial intelligence. Tertiary Education Quality and Standards Agency.

McDonald, N., Johri, A., Ali, A., & Hingle Collier, A. (2025). Generative artificial intelligence in higher education: Evidence from an analysis of institutional policies and guidelines. Computers in Human Behavior: Artificial Humans, 3, Article 100121. https://doi.org/10.1016/j.chbah.2025.100121

Miao, F., & Holmes, W. (2023). Guidance for generative AI in education and research. UNESCO. https://doi.org/10.54675/EWZM9535

Mikalef, P., & Gupta, M. (2021). Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance. Information & Management, 58(3), Article 103434. https://doi.org/10.1016/j.im.2021.103434

Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. Nature Machine Intelligence, 1, 501-507. https://doi.org/10.1038/s42256-019-0114-4

Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2020). From what to how: An initial review of publicly available AI ethics tools, methods and research to translate principles into practices. Science and Engineering Ethics, 26, 2141-2168. https://doi.org/10.1007/s11948-019-00165-5

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., & Giovannini, E. (2008). Handbook on constructing composite indicators: Methodology and user guide. OECD Publishing. https://doi.org/10.1787/9789264043466-en

National Institute of Standards and Technology. (2023). Artificial intelligence risk management framework (AI RMF 1.0) (NIST AI 100-1). https://doi.org/10.6028/NIST.AI.100-1

OECD. (2019). Recommendation of the Council on artificial intelligence (OECD/LEGAL/0449). OECD.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, n71. https://doi.org/10.1136/bmj.n71

Perkins, M., Furze, L., Roe, J., & MacVaugh, J. (2024). The AI Assessment Scale: A framework for ethical integration of generative AI in educational assessment. Journal of University Teaching & Learning Practice, 21(6). https://doi.org/10.53761/q3azde36

Plantin, J.-C., Lagoze, C., Edwards, P. N., & Sandvig, C. (2018). Infrastructure studies meet platform studies in the age of Google and Facebook. New Media & Society, 20(1), 293-310. https://doi.org/10.1177/1461444816661553

Pohle, J., & Thiel, T. (2020). Digital sovereignty. Internet Policy Review, 9(4). https://doi.org/10.14763/2020.4.1532

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, 33-44. https://doi.org/10.1145/3351095.3372873

Regan, P. M., & Jesse, J. (2019). Ethical challenges of edtech, big data and personalized learning: Twenty-first century student sorting and tracking. Ethics and Information Technology, 21, 167-179. https://doi.org/10.1007/s10676-018-9492-2

Saaty, T. L. (1980). The analytic hierarchy process. McGraw-Hill.

Saltelli, A. (2007). Composite indicators between analysis and advocacy. Social Indicators Research, 81, 65-77. https://doi.org/10.1007/s11205-006-0024-9

Selwyn, N. (2019). Should robots replace teachers? AI and the future of education. Polity.

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. American Behavioral Scientist, 57(10), 1510-1529. https://doi.org/10.1177/0002764213479366

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. Journal of Business Research, 104, 333-339. https://doi.org/10.1016/j.jbusres.2019.07.039

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., et al. (2018). PRISMA extension for scoping reviews: Checklist and explanation. Annals of Internal Medicine, 169(7), 467-473. https://doi.org/10.7326/M18-0850

UNESCO. (2021). Recommendation on the ethics of artificial intelligence. UNESCO.

van Dijck, J., Poell, T., & de Waal, M. (2018). The platform society: Public values in a connective world. Oxford University Press.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature. Organizational Research Methods, 3(1), 4-70. https://doi.org/10.1177/109442810031002

Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltynek, T., Guerrero-Dib, J., Popoola, O., Sigut, P., & Waddington, L. (2023). Testing of detection tools for AI-generated text. International Journal for Educational Integrity, 19, Article 26. https://doi.org/10.1007/s40979-023-00146-z

Williamson, B. (2017). Big data in education: The digital future of learning, policy and practice. SAGE.

Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. Learning, Media and Technology, 45(3), 223-235. https://doi.org/10.1080/17439884.2020.1798995

Wu, C., Zhang, H., & Carroll, J. M. (2024). AI governance in higher education: Case studies of guidance at Big Ten universities. Future Internet, 16(10), Article 354. https://doi.org/10.3390/fi16100354

Zawacki-Richter, O., Marin, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? International Journal of Educational Technology in Higher Education, 16, Article 39. https://doi.org/10.1186/s41239-019-0171-0