AI Beyond Prompting: Designing an Inclusive Co-Intelligence Framework for Higher Education

Kshitij Taluja
Undergraduate Researcher, Bachelor of Computer Science, Golden Gate University
Dr. Smrite Goudhaman
Global Professor of Practice, Golden Gate University

Abstract
Mollick is asking us to see AI as a partner. However, this article is asking universities an even bigger question: how can a university ensure that everyone benefits from AI, not only those able to read it, use it, and afford it? The article takes the viewpoint of resource-constrained higher education, in which language proficiency, disability status, connectivity, and cost can shape who benefits. Research shows that technology tends to accentuate existing differences rather than reduce them, and scholarship on digital inclusion says access alone does not mean similar benefits. This article presents the Inclusive Co-Intelligence Framework (ICF), developed by synthesizing literature on digital inequality, AI ethics, accessibility, and institutional preparedness. The framework applies inclusion as a cross-cutting lens across six institutional decision domains: access and provision, human-AI collaboration design, faculty readiness, student readiness, ethical safeguards, and governance and accountability. Each is rated on a continuum from absent to embedded. Assessment is therefore conjunctive: strength in one area does not offset weakness in another. This article is conceptual and its worked example is fictional; the framework is offered as a design tool and as a vehicle for future action-research cycles.
Keywords: co-intelligence; inclusive AI; higher education; AI readiness; accessibility; AI governance; resource-constrained institutions

Research questions

The purpose of this study was to develop a theoretical framework and to address conceptual questions regarding the design of an inclusive co-intelligence framework. These are therefore design questions, answered not through the collection of empirical data but through the synthesis reported in Section 5.

RQ1. How can inclusion be conceptualized at the institutional level as comparable benefit instead of access, and what does that definition require of a framework for AI adoption?

RQ2. What institutional decision-making domains must the framework cover if potential access and provision barriers (disability, language and dialect, connectivity, devices, costs) are to be identified before an AI solution is rolled out rather than after?

RQ3. How can maturity in those domains be defined so that ethical safeguards and governance functions become design requirements rather than post facto controls?

Section 5.7 explains where each question is addressed.

1. Introduction

Universities are adopting artificial intelligence faster than they can articulate how it should be implemented. Generative tools have prompted procurement, pilots, and institutional commitments, while guidance on implementation has often lagged (UNESCO, 2023, 2025). What is less often specified is whom deployment is intended to support. The question is not simply whether an institution can adopt AI, but whether the conditions of adoption allow different learners to benefit from it.

The issue is not new. While technology has the ability to increase or enhance an individual's abilities; and while the potential exists for technology to increase social inequality (Toyama, 2015); digital inclusion research continues to indicate that access to resources and technologies does not necessarily equate to similar benefits (Warschauer, 2003) – particularly in higher education settings.

There have been recent studies on the inequitable uses of generative AI in higher education, which show significant variations based upon previous experience with AI, self-efficacy/confidence, and institutional and educational support systems (Chan & Hu, 2023).

The authors define "inclusion" here as "comparable benefit" (i.e., learners derive comparable educational value or benefit from the program regardless of their level of technical skill or background). This distinction matters because even providing the same opportunities to all students may result in different educational values being derived by each student.

Much of the literature informing practice focuses on the individual. Ethan Mollick's Co-Intelligence (2024), for example, offers practical guidance for working with AI as a collaborator. This paper builds on a prior critical reading of that approach (Goudhaman & Taluja, 2026) but moves the question to the institutional level: the learner can only use the advice if the institution has made the tool reachable, usable, and supportable.

The barriers and remedies are therefore institutional. Cost, licensing, language coverage, accommodations, connectivity, and training are shaped by institutional decisions. The paper uses resource-constrained institutions as its operative category; such constraints can occur in wealthy countries as well as the Global South. The Global South is the paper's vantage point, not a synonym for resource constraint.

Existing readiness approaches provide useful ways to examine institutional capability, but they organize assessment around distinct capability areas rather than applying a comparable-benefit test across each one (Digital Education Council, 2025; EDUCAUSE, 2024; Jöhnk et al., 2021). The ICF asks about the people who will be excluded from the capability being measured by assessment tools: that is, whose ability to develop the capability is left invisible.

This paper does not argue that AI inevitably leads to increased educational inequality. Rather, this paper suggests that if the institution does not treat inclusion as a design question during its implementation of AI, then there may be an increase in educational inequality based on previous inequalities that exist in society.

Therefore, this paper provides the ICF as a way for institutions to include all learners with regard to their decision making, and as such, treats inclusion as the lens through which each institutional decision is viewed and evaluated, rather than treating inclusion as a separate evaluation or scoring system.

In addition, the ICF defines fairness within the context of education as providing equitable access and opportunity to learning resources. This means that fairness should be defined as equity (i.e., making available different levels of resources to students who have different levels of prior knowledge) as opposed to equality (e.g., providing each student the same level of resource).

Although this paper uses the university as the primary example, the logic used in developing the ICF is applicable to other types of organizations that lack sufficient resources to implement technology and/or other programs for their customers/users/learners (including non-profits and local/governmental agencies). Therefore, the ICF can help these organizations evaluate whether they are implementing fair practices regarding procurement, training, governance and access.

The purpose of this paper is to describe a new conceptual model for evaluating and enhancing the use of technology in educational settings; thus, it is a conceptual/model-building study (framework-development) rather than an empirical study.

Section 2 establishes the location of this paper at the intersection of AI ethics and AI capability. Section 3 reviews some of the most significant literature related to both issues. Section 4 describes the gap identified in the review. Section 5 explains the methodological approach to developing the ICF. Section 6 explains and illustrates the components of the ICF. Section 7 works the framework through a hypothetical case. Section 8 identifies some potential areas of concern regarding governance and safeguards associated with implementing the ICF. Section 9 outlines how educators might use the ICF to create opportunities for ongoing action-research cycles. Sections 10 and 11 identify some of the limitations of this work and present some conclusions.

2. From AI Ethics to AI Capability

A global consensus exists among nations regarding the ethics of artificial intelligence. Fairness, transparency, accountability, safety, and non-discrimination are the primary principles agreed upon internationally. These principles were documented within the UNESCO Recommendation on the Ethics of Artificial Intelligence (UNESCO, 2021), in which inclusion is defined as both a value and a principle to be implemented through policy action. UNESCO's education guidance for implementing these principles was further set out in the Beijing Consensus (UNESCO, 2019) and in guidance on generative AI in education and research (Miao & Holmes, 2023).

However, there remains an operational challenge. Evaluations of AI ethics guidance find considerable agreement on principles (Jobin et al., 2019; Mittelstadt, 2019) but significant disagreement about how those principles are to be implemented. Principles can outline what an organization should value, but they do not dictate which tools to license, which languages to support, or whose complaints to answer, nor do they provide clarity on how to distribute limited resources. The ICF addresses this implementation gap and therefore serves as guidance complementary to normative instruments rather than as a replacement for them.

In higher education, the implementation gap is made visible because ethical safeguards and capability building are typically addressed separately. Procurement, professional development, infrastructure, curriculum, and governance are each managed independently, with their own stakeholders and timeframes. It is therefore possible for an institution to develop a defensible responsible-use policy yet implement a tool that produces accessibility barriers or connectivity problems. The problem is no longer simply what values an institution holds, but whether those values are taken into account when decisions about providing AI are made.

The ICF therefore treats inclusion as the bridge between these two tracks: when an institutional value is translated into a design question with an owner, an evidence requirement, and a resource implication, the issue shifts from statements of principle to procurement, instructional design, technical support, and governance.

3. Literature Review

Individual co-intelligence. Practitioner writing on human-AI collaboration is addressed to individuals and is productive at that level. Mollick (2024) is the best example, offering four rules for working with generative systems: invite AI to the table; treat it as a person while telling it what person to be; remain the human in the loop; and assume today's system is the worst you will use. Such work tells a competent individual how to collaborate with AI and says very little about which individuals are positioned to take up the advice.

Education and AI. Research on AI in post-secondary education is large and expanding rapidly, but much of it centers on effectiveness and adoption. Zawacki-Richter et al. (2019) reviewed 146 studies of AI in higher education and reported very limited engagement with ethical aspects, risks, and challenges, and minimal involvement by educators themselves. Holmes, Bialik, and Fadel (2019) provide an applied perspective on the opportunities AI offers for teaching and learning.

Institutional readiness and maturity models. The Digital Education Council's Ten Dimension AI Readiness Framework (2025) provides ten dimensions, four guiding principles, and four readiness levels for institution-wide AI integration. EDUCAUSE's Higher Education Generative AI Readiness Assessment (2024) organizes assessment into Strategy, Governance, Technology, Workforce, and Teaching and Learning and recommends cross-functional participation. Jöhnk, Weißert, and Wyrtki (2021) derive five categories of organizational AI readiness factors—strategic alignment, resources, knowledge, culture, and data—from interviews with 25 AI experts. These models provide valuable capability structures; ICF differs in making comparable benefit the question applied across each domain rather than a separate capability result.

Neighboring literatures. Four established literatures contribute elements of ICF's logic. Universal Design for Learning argues for anticipating barriers rather than retrofitting accommodations (Meyer et al., 2014), but operates primarily at curriculum and instruction; ICF carries the design-first logic upstream to institutional adoption and procurement. Digital-divide research distinguishes access from skills and usage (Hargittai, 2002) and from differences in tangible outcomes (van Deursen & Helsper, 2015). Technology-amplification research explains how tools can reinforce existing capacities and inequalities (Toyama, 2015). Capability maturity modeling supplies the logic of assessing institutionalization across maturity levels (Paulk et al., 1993). ICF combines these strands at the point of institutional AI adoption.

Access, provision, and the gap between them. The barriers the ICF emphasizes are documented in their own literatures and largely absent from institutional adoption guidance. For disabled learners, accessible design has an auditable anchor in the Web Content Accessibility Guidelines (W3C, 2023), which give procurement a testable standard rather than a statement of intent. There is very little evidence about how disabled students experience generative AI in higher education. Based on their review of algorithmic bias in education, Baker and Hawn (2022) identify students with disabilities, international students, and students from low socio-economic backgrounds as three groups whose experiences remain poorly studied — precisely the groups an inclusion-focused framework must center. Language-related issues are best understood as differences in the quality and reliability of outputs across lower-resourced languages and dialects rather than as outright unavailability, so the framework frames language coverage as a provision issue to be validated locally at procurement rather than presumed. The access conditions that shape use — connectivity, shared devices, and the ability to afford paid tiers — are framed here as varying by institution and cohort rather than only by country; this is a design assumption rather than a confirmed distribution, proposed for testing in Section 10.

Synthesis. Three patterns emerge: individual guidance assumes conditions of access it does not examine; efficacy studies rarely report benefit separately for learners least likely to reach the tool; and institutional readiness models assess capability through parallel domains rather than a cross-domain comparable-benefit test. ICF is proposed to fill that space with an institutional, design-first method for adoption decisions.

Table 1. Neighboring frameworks compared

Framework / literature	Level of analysis	Treatment of inclusion	Operational instrument?
Universal Design for Learning (Meyer et al., 2014)	Curriculum and instruction	Design-first principle	Yes, for curriculum
Digital divide research (Warschauer, 2003; Hargittai, 2002; van Deursen & Helsper, 2015)	Individual and societal	Central; reach vs benefit	No
Technology amplification (Toyama, 2015)	Societal and program	Central; mechanism of inequality	No
AI ethics instruments (UNESCO, 2021; OECD, 2019)	International policy	Named value	No
AI readiness models (Digital Education Council, 2025; EDUCAUSE, 2024; Jöhnk et al., 2021)	Institutional	Parallel capability domains; no cross-domain comparable-benefit test	Yes, for capability
Co-intelligence (Mollick,	Individual	Assumed via access	Rules for

Framework / literature	Level of analysis	Treatment of inclusion	Operational instrument?
2024)	practice		individuals
ICF (this paper)	Institutional adoption decisions	Organizing lens across all dimensions	Yes: six dimensions with maturity levels

4. The Gap This Paper Addresses

The three gaps below are differences between what a university requires and what existing guidance provides, not errors in that guidance. No claim is made that Mollick's work was intended as institutional policy, or that readiness models were intended as equity instruments. The gaps matter because higher education has read these works as answering questions they did not set out to answer.

The first gap is the unit of analysis. Mollick (2024) wrote for the practitioner, about one person working with one system, and his guidance is appropriate at that level. Outside it sits almost every condition required to follow the advice: whether a tool is licensed, whether staff have time to learn it, whether assessment policy permits its use, whether data protection procedures must be revised. These are institutional decisions, and none appear in individual-level guidance.

The second gap, and the one this paper treats as central, is the prerequisite of access. The question raised by the prior book review (Goudhaman & Taluja, 2026) is whether evidence of within-group gains among people who already have the tool tells us who reaches it. In the field experiment by Dell'Acqua et al. (2026), consultants using AI performed better on tasks within the tool's effective range, with larger gains among lower-performing consultants; on tasks outside that range, AI users performed worse, consistent with miscalibrated trust. Noy and Zhang (2023) likewise found productivity and quality gains in a writing experiment, with larger gains among lower-performing writers. Both studies concern participants who already had access, so neither establishes how access is distributed across a wider student population.

For a framework defining inclusion as comparable benefit rather than equal access, that distinction is decisive. A tool may compress differences among its users while leaving disparities across the whole student body untouched or wider if access is concentrated among the already advantaged. The relevant institutional questions are therefore concrete: do models perform adequately in the languages students use; do interfaces work with assistive technologies; do tools function on limited bandwidth and modest hardware; and is the more capable version behind a paywall?

The third gap lies in the institutional literature that might catch the first two. Readiness and maturity models assess institutions at the right level and provide useful capability structures, but their domains are generally assessed as parallel areas. What is missing is a method that asks the inclusion question at each decision point and makes a failure in one domain visible rather than allowing strength elsewhere to obscure it. Sections 5 and 6 set out one such method.

5. Framework Development Methodology

5.1 Design

This article is a conceptual, framework-development study rather than an empirical one. As Jaakkola (2020) describes it, it belongs to theory synthesis: concepts taken from different literatures are merged to form a framework for a problem that none of them addresses alone. The method used had five phases: literature identification; selection and scope; concept-centric analysis; derivation of decision domains and dimensions; and derivation of maturity levels. Phases 1 and 2 report the purposive basis on which literature was identified and selected. Phases 3 to 5 show the analytical process by which the reviewed material was converted into the structure of the framework, and they are described in full so that the reader can audit the development of the framework rather than accept it on trust.

5.2 Literature identification

Literature identification used a targeted, concept-centric method rather than a systematized database review. Six conceptual domains (individual co-intelligence; digital inequality; accessibility; algorithmic bias in education; AI ethics; institutional AI readiness) were used to identify literature. A combination of targeted scholarly retrieval and direct access to primary policy and standards documents was used to find relevant works within those domains. If a retained document referenced a secondary document which addressed an uncovered concept, it was then selectively pursued. The goal of this process was analytical coverage of the concepts required to develop the framework rather than complete collection.

There was no time constraint placed upon the publications searched and there was also no citation tracking protocol performed either retrospectively or prospectively. Because no search log was kept during the literature search, this article does not report an exact search string, record count, or total number screened at database level; doing so would create false precision. Instead, the criteria used to select sources are set out in Section 5.3, and the analysis, synthesis, and derivation steps are described in Sections 5.4—5.6.

To test whether the purposive corpus had missed work capable of changing the framework, a verification search was carried out during revision on August 11, 2026 across publicly available scholarly search results and the websites of the publishers, standards bodies, and policy organizations cited in this article. The searches used the strings “AI readiness higher education institution framework”, “generative AI equity higher education access”, “digital divide outcomes benefit inequality”, “algorithmic bias education disability students”, “accessibility generative AI screen reader higher education”, and “AI maturity model institution capability”. Results were screened for relevance to the framework’s six conceptual domains and for evidence capable of adding, removing, or redefining a decision domain or maturity level. The search was confirmatory rather than generative and was bounded by the six domains already identified in the synthesis rather than intended as exhaustive coverage. The verification search identified additional relevant material on institutional AI readiness, maturity modeling, accessibility, equity, and higher-education AI governance, but did not identify a source requiring the six dimensions or four maturity levels to be restructured. The search details are reported so that the coverage of the synthesis can be re-tested, while the absence of a contemporaneous search log for the original purposive identification is retained as a methodological limitation in Section 10. All policy instruments referenced here (UNESCO, 2019, 2021, 2023, 2025; OECD, 2019; W3C, 2023) were consulted directly as primary documents rather than located through a database search.

5.3 Selection and scope

Selection was purposeful rather than protocol-driven. A source was retained when it addressed at least one of the six conceptual domains and contributed evidence about who benefits from technology, differential access or benefit, institutional implementation, or a mechanism relevant to one of the proposed decision domains. Material was excluded from the synthesis when its problem resided solely in individual practice or national policy, when it addressed effectiveness without relevance to differential benefit, or when it repeated a point already adequately covered by a retained source.

No PRISMA-style screening process was followed and no reliable count of records identified, screened, excluded, and retained was maintained; no such numbers are reported here rather than reconstructed after the fact. No language, source-type, or date restrictions were applied beyond the domain and relevance judgments described above. The review was concept-centric rather than exhaustive. Its intent was analytical coverage of the concepts the framework relies upon — individual co-intelligence, digital inequality, accessibility, algorithmic bias in education, AI ethics instruments, and institutional readiness — rather than comprehensive coverage of any one field. That choice is revisited as a limitation in Section 10.

5.4 Analysis and synthesis

Each retained source was read concept-centrically rather than one at a time, in the manner of the concept matrix recommended by Webster and Watson (2002). Each source was read for two things: the problem it identifies about who benefits from a technology, and the level at which that problem resides — individual practice, curriculum, institution, or policy. Problems were then compared across literatures rather than within them, and at this point recurring barriers become apparent: what is treated as a skills divide in one literature is an accessibility conformance question in another and an unexamined assumption in a third.

Three tests were applied before any recurring problem was carried forward. First, it had to appear in more than one of the literatures reviewed, so that the architecture did not rest on a single source. Second, it had to be capable of being acted upon by an institution rather than belonging to individual practice or national policy. Third, it had to bear on who benefits, not only on whether a tool functions; problems of effectiveness alone were set aside, since the framework assumes the institution has already judged the tool useful.

5.5 From recurring problems to decision domains

Problems that survived those tests were grouped together according to the institutional decision that would have to change in order to address them. The grouping of decisions rather than the grouping of themes identifies decision areas an institution can act on:

although an accessibility failure, a language gap, a device assumption, and a paywall are identified as separate issues in the literature, these issues are either resolved or entrenched at the exact moment an institution determines which tools it will specify and license, and they therefore form a single decision domain.

A cluster was retained as a dimension when four conditions were satisfied: it named a decision an institution actually makes, with an identifiable owner and budget line; the inclusion question could be asked of it meaningfully; it could not be reduced to another cluster; and maturity in it could be shown through documents and processes rather than only attitudes. Faculty and student readiness illustrate why this matters: both concern preparedness but are settled by different decisions—workload and professional development for faculty, curriculum and targeted support for students—and an institution can be strong in one and absent in the other. Ethical safeguards and governance are similarly distinct: safeguards specify the protections that should operate, while governance specifies who owns, enforces, measures, and remedies them. Six clusters met all four conditions and became the dimensions. The number is a product of the procedure rather than a claim that six is uniquely correct; Section 10 treats that choice as a limitation.

5.6 Deriving the maturity levels

The maturity levels answer a distinct question: not what an institution decides, but how deeply that decision is institutionalized. The structure follows capability maturity modeling (Paulk et al., 1993). Four levels were used rather than reproducing the full maturity sequence because the ICF measures the institutionalization of an inclusion commitment rather than process optimization. The highest state therefore combines routine budgeting, monitoring, and review rather than claiming a validated progression toward continuous optimization.

Each level was then written dimension by dimension by identifying observable evidence—documents, budget lines, audits, procedures, or records—that could place an institution at that level. Assessment is intended to be conducted by a cross-functional institutional team, with evidence checked by the relevant process owner and, where appropriate, accessibility or learner-support expertise. The levels are ordinal with respect to institutionalization only; no claim is made that they form an empirically validated growth path or that a higher level produces better learning outcomes. The resulting dimension-by-level matrix is presented in Table 3 in Section 6.

5.7 Traceability

Table 2 establishes the linkages from literature, to concept or problem, to institutional decision domain, to ICF dimension for each of the six dimensions, so that readers can assess for themselves whether each dimension arises from the material reviewed or is simply asserted.

The research questions are answered as follows: RQ1 through the definition of inclusion as comparable benefit developed in Section 1 and operationalized in Section 6; RQ2 through the derivation in Section 5.5 and Table 2; RQ3 through the maturity logic in Section 5.6 and the governance mechanisms described in Sections 8 and 9.

Table 2. Traceability from literature to ICF dimension

Literature and evidence base	Recurring problem identified	Institutional decision domain	ICF dimension
Warschauer (2003); Hargittai (2002); van Deursen & Helsper (2015); W3C (2023); UNESCO (2021)	Access does not produce benefit; disability, language, connectivity, device, and cost barriers determine who can use a tool	Accessibility conformance, language coverage, licensing, and cost model at procurement	Access and provision
Mollick (2024); Dell'Acqua et al. (2026); Toyama (2015)	Benefit depends on how tasks are structured; over-reliance arises outside a system's effective range, and tools amplify existing capacity	Task and assessment design, verification norms, and human-in-the-loop rules held in common	Human-AI collaboration design
Jöhnk et al. (2021); EDUCAUSE (2024); Digital Education Council (2025); Zawacki-Richter et al. (2019)	Institutional intent does not reach teaching without staff capability, time, and incentives; educators are largely absent from the research base	Workload models, protected time, professional development, and specialist support	Faculty readiness
Hargittai (2002); van Deursen & Helsper (2015); Chan & Hu (2023); Warschauer (2003)	Students arrive with unequal prior exposure and skill, so equal permission reproduces the second- and third-level divides	Curriculum AI literacy and targeted support for least-exposed students	Student readiness
UNESCO (2021); Jobin et al. (2019); Mittelstadt (2019); Baker & Hawn (2022); Holmes et al. (2019)	Ethical principles converge but do not specify mechanisms; bias is least studied for the groups most exposed to it	Data protection, consent, integrity rules, and bias monitoring as operating requirements	Ethical safeguards
UNESCO (2021); OECD (2019); Miao & Holmes (2023); Mittelstadt (2019)	Oversight, accountability, and remedy are asserted as norms without an owner, a redress route, or measurement	Ownership with authority, redress procedures, disaggregated measurement, and review cadence	Governance and accountability

6. The Inclusive Co-Intelligence Framework

Readiness for AI is usually evaluated by asking whether an institution can prepare itself, and less often by asking who that preparation serves. An institution may demonstrate strong capability while delivering unequal benefit across learner groups.

In ICF, inclusion is not a seventh dimension or a score added to the others. It is the cross-cutting test applied to each of the six substantive decision domains: does this domain enable comparable benefit, and if not, who is excluded and why? This makes the assessment conjunctive: strength in one domain cannot compensate for a failure that prevents comparable benefit in another. The claim is about the evaluative structure, not about the novelty of the component concepts, and it remains to be tested empirically.

The framework is written for resource-constrained institutions, where budget, connectivity, licensing capacity, or staff time can determine who is reached. Access and Provision is treated as foundational because the other domains cannot operate for a learner who cannot use the tool; Governance and Accountability is sustaining because it keeps the other five in place as tools, vendors, and staff change.

Access and Provision asks whether every learner can practically use the AI adopted by the institution: compatibility with assistive technologies such as screen readers and captioning; coverage of relevant languages and dialects; the connectivity and device standard assumed; and the cost or paywall that determines who receives the more capable version. WCAG 2.2 (W3C, 2023) gives this dimension an auditable procurement standard rather than a statement of intent. The framework was nevertheless developed without the participation of learners with disabilities, a major limitation identified in Section 10.

Human-AI Collaboration Design describes the way in which co-intelligence enters teaching through design. It also questions whether tasks maintain human judgment at their center point, whether AI operates as a collaborative supervisor or an oracle, and whether the design of the task provides safeguards against excessive reliance on AI. In a field study by Dell'Acqua et al. (2026), AI assistance resulted in decreased performance on tasks beyond the limits of its effectiveness. Therefore, this dimension at an institutional level relates to shared task design practices and verification standards, not the ability of individual instructors to apply them.

Faculty Readiness refers to the ability, training, time, incentives, and attitudes required for staff to teach with AI. Participation in a single optional workshop does not qualify as readiness; providing protected time, formalizing staff workload recognition, and providing accessible-design support provide the actionability needed from an institutionally-based commitment.

Student Readiness involves determining if students can utilize AI both ethically and critically, while acknowledging that there exist unequal starting points. Digital divides illustrate that even those who have access to technology possess varying levels of skill and usage. Hargittai (2002) illustrated the existence of the "second-level" digital divide, in which individuals may have similar levels of access yet differing amounts of use, and van Deursen and Helsper (2015) showed relationships between digital engagement and disparate tangible outcomes. Inclusive student readiness therefore requires explicit AI literacy instruction and targeted support rather than blanket approval. Student readiness is independent of faculty readiness due to differences in institutional responsibilities and decision-making authority.

Ethical Safeguards refers to the protections related to the collection and use of data regarding students, including privacy issues, obtaining informed consent, ensuring academic integrity, monitoring for biases and providing transparency concerning all aspects of using educational AI. These protections are considered inclusive when they protect the most vulnerable learners first. Baker and Hawn (2022) identify students with disabilities, international students, and students from low socio-economic backgrounds as groups that are typically underrepresented in educational AI research. Bias monitoring should therefore be treated as a design consideration, not an assertion that specific models contain biases against all members of these categories.

Governance and Accountability describes who is accountable to the AI program's stakeholders; to whom a student or teacher can address their complaints about failures of the AI system; and through what methods will institutions determine if inclusion in the use of AI systems is increasing: policy; a person designated as responsible for the program; a method students and teachers may complain through; measures to assess improvements in inclusion; and regular reviews. Ethical Safeguards define protection(s); Governance identifies who has responsibility for developing, implementing, monitoring/enforcing those protections, and remedies. Section 8 develops this mechanism.

Although the six areas overlap and relate to one another, they are not interchangeable. Access and Provision establishes a foundation upon which all other areas build; Governance and Accountability sustains that foundation; Faculty and Student Readiness represent two distinct sub-populations that require independent consideration by the institution regarding institutional decision-making processes. The organization of these six dimensions is therefore purposeful, rather than a claim that the area can be divided in only one way.

Levels of Maturity. Each of the six dimensions will be evaluated using a four-level model: Level 0, Absent: the area is ignored/does not exist. Level 1, Reactive: problems are addressed as they occur (i.e., on a case-by-case basis) with no established policy. Level 2, Structured: an established policy and allocated resources exist, but implementation may vary. Level 3, Embedded: the area has been designed into institutional infrastructure/budgeting and includes ongoing assessment/evaluation of performance. The progression of institutionalization of capabilities follows the same reasoning as Capability Maturity Modeling (Paulk et al., 1993). Evidence rather than assertion is required for placement at each level: for example, an accessibility audit and documentation of cost and language coverage for Access and Provision; documented common task-design guidance for Human-AI Collaboration Design; documentation of protected training time for Faculty Readiness; a curriculum-based AI literacy plan for Student Readiness; and both a named contact and a documented redress process for Governance and Accountability. Future work includes development of a validated scoring rubric for the entire tool.

Table 3. The ICF maturity matrix

Dimension	Level 0: Absent	Level 1: Reactive	Level 2: Structured	Level 3: Embedded
Access and provision	No accessibility, language, connectivity, or	Barriers addressed only after	Access policy exists; applied unevenly across	Accessibility, language, device, and cost checks

Dimension	Level 0: Absent	Level 1: Reactive	Level 2: Structured	Level 3: Embedded
	cost review before adoption	individual complaints	tools	are procurement conditions, audited against WCAG 2.2
Human-AI collaboration design	AI use in teaching unconsidered	Task problems addressed after incidents	Shared task-design guidance exists; uptake varies	Institution-wide task designs keep human judgment central; over-trust guarded by design
Faculty readiness	No training or support	One-off optional workshop after problems	Structured training; no protected time or incentives	Training with protected hours, workload recognition, and accessible-design support
Student readiness	Student AI capability assumed or ignored	Blanket bans or blanket permission after incidents	AI literacy optional or one-time	AI literacy taught in curriculum, with targeted support for least-exposed students
Ethical safeguards	No AI-specific data, consent, or integrity provisions	Cases handled ad hoc	Documented policy; enforcement uneven	Safeguards protect most-exposed learners first; bias monitoring routine and reviewed
Governance and accountability	No owner, no policy, no complaint route	Issues escalate informally	Named committee and policy; no redress route or measurement	Named owner with authority to delay rollout; usable learner redress; inclusion measured and reported

The contribution, stated precisely. The individual components assembled here are established: UDL supplies a design-first logic; digital-divide research distinguishes access, skills, and outcomes; capability maturity modeling supplies institutionalization levels; and AI-readiness work supplies organizational capability domains. ICF's proposed contribution is their integration at institutional AI-adoption decisions: inclusion is the cross-cutting test, comparable benefit is the outcome criterion, the domains are tied to concrete decisions, and maturity levels make institutionalization visible. The framework is therefore an operational design proposal, not a claim of conceptual priority over these literatures.

7. Operationalizing the ICF: A Constructed Illustration

A framework earns its place by demonstrating something an institution can see that it would otherwise miss. **The case presented below is constructed and hypothetical. It is neither a study, nor a pilot, nor a report of anything that happened.** No institution was observed, no student was consulted, and nothing here should be read as a finding; it is a design walkthrough developed to make the six dimensions concrete. Empirical application is proposed in Sections 9 and 10.

The case study example uses University A (a publicly funded institution), operating under the resource constraints described in Section 6. The generative AI assistant was contracted for use as a tool for undergraduate students. The case study follows S, a constructed composite of barriers documented in the literature to make questions from the framework visible. S is a first-year undergraduate student who uses a screen reader. She studies in a language other than her most fluent reading/writing language. She connects mainly through mobile data on a shared device, and she cannot afford the paid tier of the adopted tool. Each statement represents what the framework would surface; none is an observed finding.

Access and Provision. Does the tool work with her screen reader? Many institutions would not know, because the question never arose before procurement. If the interface cannot be navigated, the license is irrelevant to her; if output quality falls in her stronger language, she is pushed to work in her weaker one, the opposite of what the tool was bought to achieve. At Level 1 the university learns of this when she complains. At Level 3, compatibility is a procurement condition and no tool is signed off without it.

Human-AI Collaboration Design. Suppose the tool is usable. If the task design requires students to verify AI output against sources themselves inaccessible to her, verification becomes a second barrier rather than a safeguard. Patterns of reliance could plausibly differ, since verification costs her considerably more time than it costs her peers — a design hypothesis the framework produces for testing, not an established effect.

Faculty Readiness. The lecturers received a single training session, covering use of the tool rather than inclusive teaching with it — under the time constraints common in under-resourced institutions, a predictable outcome rather than an individual failing. A Level 3 position supplies protected time, recognition in the workload model, and someone who can answer an accessibility question quickly.

Student Readiness. S has had less prior exposure to AI tools than classmates who used them throughout school, and such exposure as she had was bounded by what her assistive technology allowed. Blanket permission therefore advantages those who arrive fluent. The gap in this construction is one of provision and prior opportunity, not of capability.

Ethical Safeguards. Interaction data generated by S may be more sensitive than other students': usage patterns associated with assistive technology could disclose her disability status to anyone with access to the logs. That is a data governance design question, not an incident to be handled afterwards. Academic integrity policy raises a parallel question — whether her accommodations and the institution's AI-use rules cohere or contradict.

Governance and Accountability. What does she do if the tool does not function for her, and what happens next? The dimension has to have a named owner, a redress route which does not rely upon her believing that she can pursue it, as well as performance indicators that will show a problem without needing her to file a complaint. Since there are no separate outcome measures for different student groups, the dashboard displays increasing use of the tool while her experience of it remains unaltered. Because the case is constructed, the maturity placements shown are illustrative and not institutional assessments. The walkthrough demonstrates how a rollout that appears healthy in aggregate could reveal weaknesses at Levels 1 or 2 across multiple dimensions, and a failure at Level 0 on Access and Provision, while a standard readiness review still concludes that the university as a whole is prepared. Therefore, this illustrates the questions ICF asks; it does not illustrate that the application of those questions results in the described outcomes. The same decision logic can be transferred, with domain-specific adaptation, to other resource-constrained public or nonprofit institutions.

8. Governance, Ethics, and Safeguards

Governance is the sixth dimension because it determines whether the other five hold over time. Access won once can be lost quietly. The grounding is the UNESCO Recommendation (UNESCO, 2021), which emphasizes human oversight, responsibility, accountability, and avoiding an increase in digital divides. OECD principles (OECD, 2019) and UNESCO guidance on generative AI in education (Miao & Holmes, 2023) point in the same direction, but none of these instruments supplies the specific institutional ownership and redress mechanism proposed here.

Ownership. Someone must bear responsibility for whether the AI program works for all learners, and that responsibility must be distinct from technical ownership of the tool. The accountable role should have authority to delay or stop a rollout when evidence shows a material inclusion failure.

Redress. The second requirement concerns redress. If there is no conventional way for a student who has experienced failure due to lack of inclusion to formally complain or escalate their issue (and therefore seek remedy), then it is the responsibility of the institution to provide some alternative means for them to do this. It is a form of governance that provides the necessary mechanisms, but does not assume that all students will advocate on their own behalf.

Measurement. In terms of measurement, while comparable benefits can only emerge through consideration of both use/usage and the outcomes of those uses relative to the level of provision provided, there is a conflict over privacy with respect to collecting information regarding a person’s disability or language status. For that reason, the framework offers several possible alternatives to collect data such as providing minimum aggregation thresholds, opt-in disclosure, and providing indicators of provision (for example, what percentage of all licensed tools meet a specified accessibility standard) so that institutions can collect comparative data relative to levels of provision without collecting more sensitive personal information than is necessary.

Review cadence. The fourth requirement concerns the frequency with which the accessibility of each tool is reviewed. Since accessibility is never static and since vendors can change either the model or interface of tools they sell (in other words, once inclusion has been established, there is no guarantee that the vendor will continue to support the same level of inclusion from one generation of product to the next), institutions require a schedule for reviewing accessibility and checking versions, and also require contractual notification prior to making any material change to the product.

9. Supporting Action Research with the ICF

This article does not report a completed action research study. No institution or other stakeholder took part in developing the framework; no change was implemented on the basis of it; and no results of its use have been reported. The ICF is a design artifact, and the claims made for it are conceptual. This section sets out how the framework is intended to be used by institutions and researchers who do undertake such work: as the instrument around which a cycle of diagnosis, action, evaluation, reflection, and adaptation can be organized.

Diagnosis. A cross-functional institutional team assesses each dimension against the evidence anchors in Table 3, using documents, audit reports, procurement records, and relevant learner-support evidence rather than perception alone. A conjunctive rule applies: a dimension is rated at the level its weakest evidence supports, and strength elsewhere never raises it.

Action. Diagnosis identifies a single binding constraint rather than a list of recommendations. Because Access and Provision is foundational, a Level 0 or 1 rating there will typically take precedence over all other areas, since interventions elsewhere cannot reach a learner who has no means of accessing the tool.

Evaluation. Change is assessed against comparable benefit rather than adoption. That requires outcome indicators disaggregated by the categories the framework names, and provision indicators that measure what the institution actually supplies rather than what students report. The tension between disaggregation and privacy discussed in Section 8 is a live design problem at this stage, not a technicality.

Reflection. Findings are interpreted together with the people the framework is about; a cycle in which disabled and multilingual learners are measured but not consulted reproduces the limitation acknowledged in Section 10. Reflection is also where an institution examines whether the dimension it treated as binding really was the limiting factor.

Adaptation. Two kinds of revision occur. The institution revises its practice and re-enters the cycle at the next dimension; the framework itself is also revised, since whether six dimensions are the right six, whether the level descriptors discriminate usefully in that setting, and whether the evidence anchors were obtainable are findings to report back rather than absorb silently. The ICF is offered as a revisable instrument, and use of this kind is how it would be revised.

Table 4. Using the ICF across an action-research cycle

Cycle stage	What the institution does	What the ICF supplies	Evidence produced

Cycle stage	What the institution does	What the ICF supplies	Evidence produced
Diagnosis	Rates six dimensions against evidence anchors, disaggregated by learner group	Dimensions, levels, conjunctive rating rule	Baseline profile with an identified binding constraint
Action	Changes a specific institutional decision, prioritizing foundational dimensions	Decision domains locating where change must occur	Documented change to procurement, workload, curriculum, or policy
Evaluation	Measures comparable benefit and provision, not adoption alone	Comparable-benefit test and provision-side indicators	Disaggregated outcome and provision data
Reflection	Interprets findings with affected learners and staff	Inclusion question applied to each dimension	Participatory account of what changed and for whom
Adaptation	Revises practice and reports back on the framework	Revisable structure of dimensions, levels, and anchors	Next-cycle plan and proposed framework revisions

10. Limitations and Future Research

The contribution is a design artifact rather than a research finding, and several limitations follow.

The ICF has not been applied in an institutional setting. The maturity levels are a proposed progression of institutionalization, not an empirically validated scale; the article does not establish that institutions move through them sequentially or that higher levels produce better learning outcomes. Section 7 is a constructed walkthrough, not evidence that the dimensions relate in practice. The concept-centric review was neither systematic nor exhaustive, no contemporaneous search log was kept for the original purposive identification, and the review draws substantially on literature and instruments developed in well-resourced settings. Empirical work in contrasting resource environments is therefore necessary.

The most consequential limitation concerns disability. Disability inclusion is central to this framework: it is the driving force behind the Access and Provision dimension, it ties that dimension to conformance with WCAG, and it provides much of the illustration in Section 7. Yet no learner with a disability was involved in developing the framework. The dimension descriptors, the level anchors, and the constructed persona in Section 7 were written about disabled learners rather than with them, which is precisely the relationship the article criticizes elsewhere. That gap restricts how far the framework can be relied upon to detect barriers: the barriers it describes are those identified in the literature and in accessibility standards, and nothing here indicates that they are the barriers disabled learners themselves would identify as most critical, or that the four levels depict a progression they would recognize. Participatory validation with disabled learners, and with multilingual learners, is therefore treated here not as one item on a future research agenda but as the necessary next stage of the framework's development, and as a precondition for using the Access and Provision descriptors as an assessment instrument rather than as a design prompt. Such validation should be collaborative inquiry, in which affected learners define the questions and the criteria rather than responding to instruments already fixed. Six dimensions were derived by the procedure in Section 5.5, which constrains but does not determine the number; another analyst could divide the same material differently. The paper also does not provide a validated scoring rubric: it specifies evidence anchors but does not calibrate them. Finally, the central claim that individual-level co-intelligence can create unequal access or benefit is treated as a risk derived from existing literature, not as a causal finding of this study. Future work should test the conjunctive assessment rule against existing readiness approaches and determine whether it changes the distribution of benefit.

Four lines of further research follow. First and most urgent, participatory validation with disabled and multilingual learners on the terms described above. Second, pilot studies in at least two universities located in contrasting resource environments, structured as the action-research cycles described in Section 9, and documenting the access conditions treated as a design assumption in Section 3. Third, a test of the framework's central and least supported claim: whether inclusion-led adoption generates a different distribution of benefit than adoption guided by parallel readiness dimensions. Fourth, long-term study of whether an embedded position persists through changes of tool, vendor, and staff.

11. Conclusion

Mollick's invitation is to bring artificial intelligence into the room. This paper begins from the question left outside that invitation: who can reach the tool, and what happens after they arrive? For a university, that question is institutional because procurement, licensing, training, support, and governance determine the conditions under which students can benefit.

Three gaps make the question easy to miss: individual guidance assumes institutional conditions it does not examine; efficacy studies establish what AI does for people who receive it rather than how access is distributed; and readiness models provide useful capability domains without making comparable benefit the test applied across each domain.

ICF responds by changing the institutional question from "are we ready?" to "ready for whom?", and by asking it of every decision domain. In this regard, the article does not offer a list of new concepts but an attempt to combine four components — comparable benefit, institutional decision-ownership, inclusion, and maturity — in a single conjunctive structure. Accessibility, for example, would become a procurement condition; inclusion would have an accountable owner; redress would exist even where there is no formal complaint or appeal process; and measurement would cover both the extent and the nature of the provision made for the learner. The underlying logic may transfer from universities to other public and non-profit organizations that also work under resource constraints.

This document presents a design proposal rather than an instrument. The ICF has not yet been applied in an institution, and the learners it centers took no part in developing it.

Participatory validation, pilot studies with participating institutions, and comparative evaluation against existing readiness models are therefore required before the framework is used as an assessment tool.

References

Baker, R. S., & Hawn, A. (2022). Algorithmic bias in education. *International Journal of Artificial Intelligence in Education, 32*(4), 1052–1092. https://doi.org/10.1007/s40593-021-00285-9

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

Dell'Acqua, F., McFowland, E., III, Mollick, E. R., Lifshitz, H., Kellogg, K. C., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2026). Navigating the jagged technological frontier: Field experimental evidence of the effects of artificial intelligence on knowledge worker productivity and quality. *Organization Science, 37*(2), 403–423. https://doi.org/10.1287/orsc.2025.21838

Digital Education Council. (2025). *Ten dimension AI readiness framework*. https://www.digitaleducationcouncil.com/resource-library-items/ten-dimension-ai-readiness-framework

EDUCAUSE. (2024). *Higher education generative AI readiness assessment*. https://library.educause.edu/resources/2024/4/higher-education-generative-ai-readiness-assessment

Goudhaman, S., & Taluja, K. (2026, July 28). *Co-intelligence: Living and working with AI* [Book review]. European International University Publishing. https://publishing.eiu.ac/research-article/co-intelligence-living-and-working-with-ai/

Hargittai, E. (2002). Second-level digital divide: Differences in people's online skills. *First Monday, 7*(4). https://doi.org/10.5210/fm.v7i4.942

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.

Jaakkola, E. (2020). Designing conceptual articles: Four approaches. *AMS Review, 10*(1–2), 18–26. https://doi.org/10.1007/s13162-020-00161-0

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 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 organisational AI readiness factors. *Business & Information Systems Engineering, 63*(1), 5–20. https://doi.org/10.1007/s12599-020-00676-7

Meyer, A., Rose, D. H., & Gordon, D. (2014). *Universal design for learning: Theory and practice*. CAST Professional Publishing.

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

Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence, 1*(11), 501–507. https://doi.org/10.1038/s42256-019-0114-4

Mollick, E. (2024). *Co-intelligence: Living and working with AI*. Portfolio/Penguin.

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science, 381*(6654), 187–192. https://doi.org/10.1126/science.adh2586
OECD. (2019). *Recommendation of the Council on Artificial Intelligence* (OECD/LEGAL/0449). https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449
Paulk, M. C., Curtis, B., Chrissis, M. B., & Weber, C. V. (1993). *Capability maturity model for software (Version 1.1)*. Software Engineering Institute, CMU/SEI-93-TR-024. https://doi.org/10.1184/R1/6575027.v1
Toyama, K. (2015). *Geek heresy: Rescuing social change from the cult of technology*. PublicAffairs.
UNESCO. (2019). *Beijing Consensus on Artificial Intelligence and Education*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000368303
UNESCO. (2021). *Recommendation on the Ethics of Artificial Intelligence*. UNESCO. https://www.unesco.org/en/legal-affairs/recommendation-ethics-artificial-intelligence
UNESCO. (2023, September). *UNESCO survey: Less than 10% of schools and universities have formal guidance on AI*. UNESCO. https://www.unesco.org/en/articles/unesco-survey-less-10-schools-and-universities-have-formal-guidance-ai
UNESCO. (2025, September). *UNESCO survey: Two-thirds of higher education institutions have or are developing guidance on AI use*. UNESCO. https://www.unesco.org/en/articles/unesco-survey-two-thirds-higher-education-institutions-have-or-are-developing-guidance-ai-use
van Deursen, A. J. A. M., & Helsper, E. J. (2015). The third-level digital divide: Who benefits most from being online? In L. Robinson, S. R. Cotten, J. Schulz, T. M. Hale, & A. Williams (Eds.), *Communication and information technologies annual* (Studies in Media and Communications, Vol. 10, pp. 29–52). Emerald. https://doi.org/10.1108/S2050-206020150000010002
W3C. (2023). *Web Content Accessibility Guidelines (WCAG) 2.2*. W3C Recommendation. https://www.w3.org/TR/WCAG22/
Warschauer, M. (2003). *Technology and social inclusion: Rethinking the digital divide*. MIT Press.
Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly, 26*(2), xiii–xxiii.
Zawacki-Richter, O., Marín, 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