

Article

Toward Sustainable AI Literacy in Higher Education: A Delphi–AHP Approach to Identifying Key Components of AI/Data Extracurricular Programs

Jung-Sup Bae ¹, Kyoung-Min Ko ^{1,*} and Hye-Jin Park ^{2,*}

¹ Division of Liberal Arts, KU Glocal Innovation College, Konkuk University Glocal Campus, Chungju 27478, Republic of Korea; bluelion10@kku.ac.kr

² Department of Education, Graduate School of Education, Konkuk University Glocal Campus, Chungju 27478, Republic of Korea

* Correspondence: drewhoney@kku.ac.kr (K.-M.K.); phj4858@kku.ac.kr (H.-J.P.)

Abstract

Artificial intelligence (AI) literacy has emerged as a foundational competency for sustainable higher education aligned with Sustainable Development Goal 4 (Quality Education). Although AI literacy coursework is expanding rapidly, classroom instruction alone is insufficient for cultivating the applied, ethical, and reflective dimensions of AI competency, and systematic research on the components of curriculum-linked AI/data extracurricular programs remains scarce. This study employed a sequential mixed-methods design integrating the Delphi technique and the Analytic Hierarchy Process (AHP) to identify and prioritize the key components of such programs. Ten experts spanning AI education, curriculum development, extracurricular administration, and industry participated in two Delphi rounds, followed by AHP pairwise comparisons and content validity ratio (CVR). The Delphi rounds yielded a consensus framework of 24 components across seven domains; six domains served as AHP Level 1 evaluation criteria and the seventh (program type) as Level 2 alternatives. AI practice integration and curricular linkage jointly accounted for 57.2% of the criterion weight, and three program types—AI tool workshops, AI-based problem-solving challenges, and data analytics projects—exhibited distinct strategic priorities. The final framework achieved a content validity index of 0.84, providing an empirically validated design framework for sustainable AI literacy education, with potential applicability to resource-constrained settings.

Keywords: AI literacy; sustainable higher education; SDG 4; extracurricular programs; Delphi technique; Analytic Hierarchy Process; curriculum-linked learning; AI competency



Academic Editor: Antonio P. Gutierrez de Blume

Received: 2 July 2026

Revised: 21 July 2026

Accepted: 23 July 2026

Published: 7 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Artificial intelligence (AI) has become deeply embedded in industry, healthcare, education, and culture, and AI literacy—the capacity to critically evaluate, communicate with, and meaningfully use AI tools—has accordingly been positioned as a universal competency required of all citizens, regardless of major or professional background [1,2]. In higher education, this transition has accelerated. Chee, Ahn, and Lee [3], through a systematic review, demonstrated that university-level AI literacy goes beyond the foundational AI knowledge typical of K-12 settings and requires deeper engagement with data and algorithms, problem-solving capacities, and career-relevant competencies. UNESCO's AI Competency Framework [4] articulates a developmental pathway from Understand

to Apply to Create, emphasizing that students must mature into responsible users and co-creators of AI technologies.

These global developments resonate with the broader sustainability agenda. The 2030 Agenda for Sustainable Development, particularly Sustainable Development Goal 4 (Quality Education), calls for inclusive, equitable, and lifelong learning opportunities that prepare learners for emerging societal and technological challenges [5]. Within this agenda, AI literacy has been identified as a critical lever for both individual employability and the long-term sustainability of higher education systems [6,7]. The World Economic Forum's Future of Jobs Report 2025 noted that 85% of employers regard upskilling their workforce in AI as a top priority [8].

However, scholars have repeatedly emphasized that classroom instruction alone is insufficient for cultivating AI literacy as a multidimensional competency. Ng et al. [9] distinguished between knowing and applying, arguing that AI literacy must encompass practical application and critical evaluation capacities, not merely declarative knowledge. Chiu et al. [10] further developed the construct of AI competency as a self-reflective, confidence-bearing capacity to apply knowledge in beneficial ways. From the lens of Kolb's experiential learning theory [11], formal coursework primarily addresses abstract conceptualization and reflective observation, while concrete experience and active experimentation—the other half of the learning cycle—require additional instructional mechanisms. Extracurricular programs, defined as university educational activities that operate without academic credit, have long been recognized as one such mechanism [12,13].

Recent scholarship has reframed extracurricular activities as co-curricular—intentionally aligned with formal curriculum learning outcomes [14,15]. Empirical studies have shown that systematically linking coursework with co-curricular programs significantly improves students' knowledge application, logical thinking, and creative problem-solving capacities [16,17]. Yet, in the specific context of AI/data education in higher education, models that systematically connect coursework with extracurricular learning remain rare, and the constituent components of such programs have not been empirically identified through rigorous decision-making methodologies [18,19].

This gap is particularly consequential given the policy environment surrounding AI education in higher education. National-level initiatives in numerous countries—including U.S. EDUCAUSE frameworks [20], European Union AI literacy strategies [21], and East Asian AI-focused university programs—are actively investing in AI education infrastructure and curricular reform. The sustainable implementation of these national-level investments depends on universities' ability to design AI/data extracurricular programs that are evidence-based, systematically integrated with coursework, and adaptable to diverse institutional contexts, including resource-constrained settings.

To address this gap, the present study employs the Delphi technique combined with the Analytic Hierarchy Process (AHP)—a methodologically rigorous mixed-method design widely applied in curriculum development and program design research [22–24]. Specifically, the study addresses three research questions:

RQ1: What are the essential components of AI/data extracurricular programs that effectively complement formal coursework in higher education?

RQ2: Among the design domains identified through expert consensus, which carry the greatest relative weight?

RQ3: Among general program types, which should be prioritized to achieve sustainable AI literacy development in higher education contexts?

By providing an empirically grounded component framework with full methodological continuity between the qualitative and quantitative phases, this study contributes to the design science of sustainable AI literacy education in alignment with SDG 4.

2. Theoretical Background

2.1. AI Literacy and Sustainable Higher Education

AI literacy is a recently emerging construct whose definitional contours continue to evolve. Long and Magerko [1] defined AI literacy as “a set of competencies that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace.” This three-dimensional formulation—technical understanding, critical evaluation, and practical application—has served as a foundational reference for subsequent scholarship.

Ng et al. [9] proposed a four-domain framework comprising knowing and understanding AI, using and applying AI, evaluating and creating AI, and AI ethics. The authors emphasized that AI literacy is a multidimensional competency extending beyond technical knowledge to encompass ethical judgment and creative application. Subsequently, Ng et al. [25] developed and validated a multidimensional AI literacy questionnaire integrating affective, behavioral, cognitive, and ethical approaches. Chiu et al. [10] introduced the broader construct of AI competency, distinguishing it from AI literacy by defining competency as the confident application of knowledge in beneficial ways. Their five-component framework—technologies, impact, ethics, collaboration, and self-reflection—incorporates a self-reflective mindset and self-efficacy as integral elements.

From the standpoint of sustainable higher education, AI literacy has been increasingly framed as a precondition for lifelong learning and adaptive citizenship in a rapidly evolving technological environment [6,7,26]. UNESCO’s AI Competency Framework for Teachers [4] explicitly positions AI literacy within the broader mission of education for sustainable development, articulating a three-tier progression—Understand, Apply, Create—structured across four competency domains: human-centered mindset, AI ethics, AI techniques and applications, and AI system design. The EDUCAUSE AI Literacy in Teaching and Learning (ALTL) framework [20] further extends this perspective to higher education, identifying technical understanding, evaluation, practical application, and ethical consideration as four core competency areas for sustainable AI integration in universities.

These frameworks share a common emphasis: sustainable AI literacy is not adequately developed through theoretical instruction alone. Effective competency formation requires authentic engagement with AI tools, opportunities for critical evaluation, and reflective practice across diverse contexts. This requirement converges with longstanding insights from the field of experiential learning regarding the inadequacy of decontextualized classroom instruction for higher-order competency development.

2.2. Theoretical Foundations of Extracurricular and Co-Curricular Programs

Extracurricular programs in higher education refer to educational activities conducted outside of formal credit-bearing coursework. Bartkus, Nemelka, Nemelka, and Gardner [27], through a comprehensive literature review, distinguished extra-curricular activities (those proceeding independently of academic coursework) from co-curricular activities (those intentionally aligned with the learning outcomes of formal curriculum). This distinction has become increasingly salient in the context of competency-based higher education, where universities seek to maximize the integrated development of students through purposeful coordination of curricular and extracurricular experiences [14,15].

Kolb’s experiential learning theory [11] provides one of the most influential theoretical foundations for curriculum–extracurricular linkage. According to this theory, effective learning proceeds through a four-stage cycle: concrete experience, reflective observation, abstract conceptualization, and active experimentation. Classroom instruction typically engages the latter two stages (abstract conceptualization and reflective observation), while well-designed extracurricular programs provide opportunities for concrete experience and

active experimentation. Kuh [14] extended Kolb's framework to higher education through the concept of high-impact educational practices, emphasizing that meaningful learning requires students to integrate experiences across curricular and extracurricular contexts and to apply learning across diverse situations.

Deci and Ryan's self-determination theory (SDT) [28,29] provides a complementary motivational framework for understanding extracurricular participation. According to SDT, intrinsic motivation is fostered by the satisfaction of three basic psychological needs: autonomy, competence, and relatedness. Because extracurricular programs depend on voluntary participation, their design must address these needs to sustain engagement. For AI/data programs specifically, this implies that students should be able to select topics of personal interest (autonomy), experience tangible skill development (competence), and engage with peers and instructors in meaningful interaction (relatedness).

Empirical research has documented the positive effects of extracurricular participation in higher education. A meta-analysis of Korean studies up to 2017 [12] reported an overall effect size of 0.534 (medium magnitude), with particularly strong effects on community-mindedness, foundational competencies, and college life adaptation. More recent studies on curriculum-linked extracurricular programs have shown significant improvements in knowledge application, logical thinking, creative problem-solving, and ethical competencies [16,17]. Importantly, these studies converge on the finding that intentional alignment between coursework and extracurricular activities is a critical determinant of educational impact.

2.3. AI/Data Education and the Need for Curriculum-Linked Extracurricular Programs

Despite the rapid expansion of AI literacy coursework in higher education institutions worldwide, the design and operation of complementary extracurricular programs in the AI/data domain remain underdeveloped both in practice and in scholarship. Most existing research on AI education in higher education has focused on conceptual frameworks for AI literacy [1,9,10], measurement instrument development [25], or pedagogical strategies for AI in classroom contexts [30]. The systematic design of curriculum-linked extracurricular programs—particularly through rigorous decision-making methodologies—has received scant attention. Recent evidence syntheses corroborate this gap: scoping and systematic reviews of AI literacy research report that existing scholarship concentrates on formal instruction, conceptualization of the construct, and assessment development, while the systematic design of extracurricular or co-curricular AI literacy programs has attracted little attention [18,19].

This gap is significant for several reasons. First, AI/data competency formation inherently requires hands-on engagement with tools, real datasets, and authentic problems. Unlike conceptual subjects, AI literacy cannot be fully developed through reading and discussion alone. Second, the rapid evolution of AI tools (e.g., generative AI platforms) demands flexible learning environments that can adapt more quickly than formal curriculum revision cycles permit. Extracurricular programs, with their flexible operational structure, are well-suited to address this need. Third, sustainable AI education in higher education requires alignment with the equity dimension of SDG 4: resource-constrained institutions must be able to participate meaningfully in AI literacy development without prohibitive infrastructure investments, a goal that extracurricular programs—especially those leveraging cloud-based and generative AI tools—can directly support. Addressing this gap requires a methodology capable of both eliciting program components inductively from expert consensus and quantifying their relative importance for design decision-making; the following section reviews the Delphi-AHP approach adopted for this purpose.

2.4. Delphi–AHP Methodology in Educational Program Design

The Delphi technique is a structured communication method that elicits expert consensus through iterative, anonymous questionnaires with controlled feedback [31]. Originally developed at the RAND Corporation in the 1950s, the technique has been widely applied in educational research for curriculum development, competency framework design, and assessment indicator construction [22,32]. The technique's principal strength lies in its capacity to aggregate expert judgment while mitigating the social–psychological biases (e.g., groupthink, dominance effects) that can compromise face-to-face deliberation.

The Analytic Hierarchy Process (AHP), developed by Saaty [23], is a multi-criteria decision-making technique that decomposes complex problems into hierarchical structures of goal, criteria, and alternatives, and then quantifies relative importance through pairwise comparisons. AHP enables the conversion of qualitative expert judgments into numerical weights, with built-in consistency verification through the Consistency Ratio (CR). In educational research, AHP has been applied to curriculum prioritization [24], competency model development [33], and assessment indicator weighting.

The combined Delphi–AHP approach is particularly well-suited to educational program design contexts where (a) existing frameworks are insufficient and components must be inductively elicited from expert consensus and (b) the relative importance of components must be quantified to inform resource allocation and design priorities [24,33]. A methodologically important feature of well-designed Delphi–AHP studies is the direct continuity between phases: the consensus domains identified through Delphi should serve as the structural foundation for the AHP hierarchy, ensuring that quantitative prioritization is grounded in qualitative consensus rather than being driven by externally imposed criteria. Recent applications—including digital health competency [33] and ESL curriculum prioritization [24]—have demonstrated the methodology's value for generating both theoretically grounded and practically actionable design frameworks.

3. Methods

3.1. Research Design

This study employed a sequential mixed-methods design [34] integrating four phases: (1) literature review and international benchmarking to establish the theoretical foundation and an initial pool of candidate components; (2) two-round Delphi survey with an expert panel to refine and validate components across seven design domains; (3) AHP analysis structured to inherit directly from the Delphi consensus, using six of these domains as Level 1 evaluation criteria and the seventh (program type) as Level 2 alternatives; and (4) content validity ratio (CVR) verification to confirm the validity of the final framework. This sequential design allows qualitative expert insight (Delphi open-ended responses) to inform quantitative prioritization (AHP pairwise comparisons), with both feeding into a final validation step (CVR). The combined design has been recommended for educational program development contexts where existing frameworks are insufficient and where actionable priority rankings are required [23,24,33].

3.2. Expert Panel Composition

The study was conducted in the Republic of Korea, a context of particular informative value for international higher education research: strong national policy investment in AI education coexists with pronounced demographic decline in the school-age population, placing non-metropolitan universities under acute resource constraints. Korean higher education thus constitutes a leading-edge case in which the tension between ambitious AI literacy agendas and limited institutional resources—a tension increasingly faced by universities worldwide—is already highly visible. Ten experts were purposively selected

to participate in all phases of the study. To strengthen the generalizability of the findings beyond a single institutional context and to ensure that the derived components reflect broad consensus across the higher education sector, panel members were drawn from diverse institutional affiliations and professional roles. Panel composition adhered to Green's [22] recommendation that Delphi panels comprise individuals with substantive expertise on the topic, and to the suggested minimum panel size for educational Delphi studies (8–15 experts) [24,33].

The final panel included experts from four domains: (a) AI/data literacy education at universities ($n = 3$, drawn from different institutions); (b) curriculum development and educational technology ($n = 3$, drawn from different universities including both metropolitan and non-metropolitan institutions); (c) university extracurricular program administration ($n = 2$, from teaching and learning centers at distinct universities); and (d) AI industry practitioners with prior education or training experience ($n = 2$). In total, the ten experts represented nine distinct organizations (eight universities—four metropolitan and four non-metropolitan—and one company), with each of the eight academically affiliated experts drawn from a different university. All panel members held a master's degree or higher (eight held doctoral degrees), and all had a minimum of three years of relevant professional experience. The deliberate inclusion of experts affiliated with non-metropolitan institutions ($n = 4$)—defined here as universities located outside the Seoul Capital Area (Seoul, Incheon, and Gyeonggi Province), commonly referred to as regional universities—was intended to ensure that the resulting framework reflects the diverse institutional realities of contemporary higher education, including resource-constrained contexts where AI infrastructure may be limited. While the study originated within a specific institutional context, the diversity of the expert panel supports the interpretation of findings as reflecting broader sectoral consensus within Korean higher education rather than institution-specific consensus; the applicability of the framework beyond this national context is addressed with caution in the Discussion (Section 5) and Limitations presented in the Conclusions (Section 6).

3.3. Delphi Survey Procedure

Two Delphi rounds were conducted between May and June 2026. Round 1 employed a semi-structured open-ended questionnaire across seven design domains: program type, learning objectives, mode of operation, curricular linkage, AI practice integration, assessment systems, and support systems. Panel members were provided with an initial pool of candidate components derived from the literature review and international benchmarking, but they were encouraged to propose additional or alternative components freely. The international benchmarking phase reviewed publicly available program documentation from twelve institutions and organizations, including the EDUCAUSE AI literacy frameworks, European Union DigComp-related documents, and AI extracurricular program materials from East Asian universities such as the National University of Singapore and the Hong Kong University of Science and Technology; sources were selected on the basis of (a) an explicit AI/data literacy focus, (b) documented extracurricular or co-curricular delivery, and (c) availability of program-level documentation sufficient for component extraction. Responses were analyzed through content analysis, with similar responses consolidated into discrete component items. The first author and one of the corresponding authors independently coded the open-ended responses and consolidated semantically equivalent items; coding discrepancies were reviewed and adjudicated by the other corresponding author.

Round 2 used a structured closed-ended questionnaire in which panel members rated each component derived from Round 1 on a 5-point Likert scale (1 = strongly inappropriate; 5 = strongly appropriate) for both relevance and importance. Feedback from Round 1

(mean, standard deviation, and interquartile range for each item) was provided to enable participants to reconsider their judgments in light of group tendencies. Consensus was determined using three indicators: (a) content validity ratio (CVR) [35] with a minimum threshold of 0.62 for a panel of 10 (i.e., at least nine panel members rating the item ≥ 4); (b) coefficient of variation ($CV \leq 0.5$); and (c) degree of consensus (≥ 0.75 based on the interquartile range to upper quartile ratio). Items meeting all three thresholds were retained; items failing to meet thresholds were eliminated. Response rates for both rounds were 100%.

3.4. AHP Procedure

A key methodological feature of this study is the direct continuity between the Delphi and AHP phases. Following the methodological logic of Tang [24] and Yao et al. [33], the AHP hierarchy was constructed by drawing all evaluation criteria and alternatives directly from the Delphi consensus, rather than introducing externally derived criteria. Specifically, six of the seven Delphi-consensus domains (learning objectives, mode of operation, curricular linkage, AI practice integration, assessment system, and support system) were mapped to Level 1 evaluation criteria, while the seventh domain (program type) was mapped to Level 2 alternatives, as program types constitute design choices rather than evaluative criteria. This procedure ensured complete methodological continuity between the qualitative consensus-building phase (Delphi) and the quantitative prioritization phase (AHP), allowing the AHP analysis to function as a direct quantification of the Delphi consensus rather than as an independent analytical step.

The resulting hierarchy was structured in three levels (Figure 1): Level 0 (overall goal: optimal composition of AI/data curriculum-linked extracurricular programs); Level 1 (six evaluation criteria: C1 learning objectives, C2 mode of operation, C3 curricular linkage, C4 AI practice integration, C5 assessment system, C6 support system); and Level 2 (three program types derived from the Delphi consensus: A1 AI tool workshops, A2 AI-based problem-solving challenges, A3 data analytics projects). The number of Level 1 criteria ($n = 6$) remained within Saaty's [23] recommended range (7 ± 2) for cognitively tractable pairwise comparison. All AHP computations—including geometric mean aggregation, eigenvector-based priority weight derivation, and consistency ratio calculation—were performed using Microsoft Excel (Microsoft 365, Microsoft Corporation, Redmond, WA, USA).

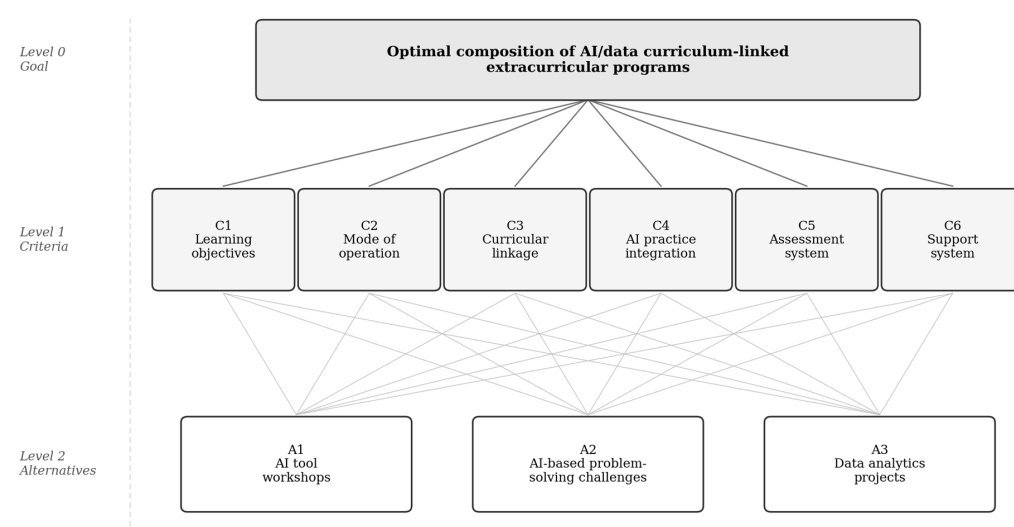


Figure 1. AHP hierarchy for prioritizing the components of AI/data curriculum-linked extracurricular programs.

Pairwise comparison questionnaires were administered to the same expert panel using Saaty's 9-point scale (1 = equal importance; 9 = absolute importance) [23]. Geometric mean aggregation was used to combine individual judgments [23,36], and the eigenvector method was applied to compute priority weights. The consistency ratio (CR) was calculated for each comparison matrix, with $CR \leq 0.1$ set as the threshold for acceptable consistency. Responses exceeding this threshold prompted requests for revision; all aggregated matrices in the final analysis met the threshold.

3.5. Content Validity Verification

The final integrated framework—comprising 24 consensus components organized into seven domains, plus the AHP priority architecture—was subjected to expert content validity verification using Lawshe's [35] CVR. The same expert panel rated each component as essential (3), useful but not essential (2), or not necessary (1). CVR was computed as $(N_e - N/2) / (N/2)$, where N_e is the number of experts rating the item as essential and N is the total panel size. For $N = 10$, the minimum acceptable CVR was 0.62 [35,37]. The content validity index (CVI) was computed as the mean CVR across all retained items, with $CVI \geq 0.80$ set as the threshold for overall framework validity [38]. Items not meeting the CVR threshold were revised based on expert qualitative feedback and re-verified through email-based confirmation. This procedure follows established practice in instrument content validation, in which expert panels iteratively assess and refine items until validity thresholds are met [35,38,39]. In addition to the quantitative CVR criterion, face validity was addressed qualitatively: panel members reviewed the wording, clarity, and surface appropriateness of each item and provided open-ended feedback, which informed the item revisions reported in Section 4.4.

3.6. Validity and Reliability

Several procedures were implemented to enhance the validity and reliability of the findings. First, expert panel diversity—across institutional affiliation, geographic location, and professional role—was deliberately ensured to mitigate single-source bias and strengthen the generalizability of the resulting framework. Second, methodological triangulation was achieved by combining literature-based theoretical analysis, international institutional benchmarking, Delphi consensus, AHP quantification, and CVR verification. Third, the direct continuity between Delphi and AHP phases—using Delphi-consensus domains as AHP Level 1 criteria—eliminated the methodological risk of introducing externally derived or post-hoc criteria into the quantitative prioritization. Fourth, AHP consistency was strictly enforced ($CR \leq 0.1$) to exclude logically inconsistent judgments. Fifth, Delphi anonymity was maintained throughout to mitigate social influence effects. Sixth, all phases were documented in an audit trail to support traceability and replicability.

4. Results

4.1. Delphi Round 1: Identification of Candidate Components

Round 1 of the Delphi survey, conducted with all ten expert panel members (100% response rate), yielded an initial set of 33 candidate components across seven domains. Notable items achieving full consensus (mentioned by all 10 panel members) included “AI tool practical application competency,” “generative AI tool utilization,” and “weekly curriculum–extracurricular topic mapping.” These findings indicate strong expert consensus regarding two foundational design principles: the centrality of generative AI in contemporary AI/data education and the necessity of systematic curriculum–extracurricular alignment.

All seven domains generated at least three candidate items, suggesting that the design of AI/data extracurricular programs is intrinsically multidimensional. Items with lower

frequency of mention (e.g., “individual self-directed learning mode” and “course credit bonus for extracurricular participation”) emerged as potentially contentious and were flagged for closer scrutiny in Round 2.

4.2. Delphi Round 2: Consensus on Final Components

Round 2 evaluated all 33 candidate components against the three consensus criteria ($CVR \geq 0.62$, $CV \leq 0.5$, $consensus \geq 0.75$). Of these, 24 components met all three thresholds and were retained, while nine items were eliminated. The nine eliminated items, together with their complete consensus statistics (mean, CVR, CV, and degree of consensus), are reported in Supplementary Table S1 to ensure full transparency of the consensus process. The final 24 components were organized into seven domains, as shown in Table 1.

Table 1. Final consensus components from the Delphi survey (N = 10).

Domain	Component	M	CVR
D1. Program type	AI tool workshops	4.70	1.00
	Data analytics projects	4.50	1.00
	AI-based problem-solving challenges	4.60	1.00
D2. Learning objectives	AI tool practical application competency	4.90	1.00
	Data interpretation and visualization	4.50	1.00
	AI ethical judgment	4.10	0.80
	Discipline-specific AI integration competency	4.40	1.00
D3. Mode of operation	Problem-solving and critical thinking	4.70	1.00
	Regular semester-based operation (curriculum-linked)	4.80	1.00
	Blended (online-offline hybrid) format	4.40	1.00
	Team-based collaborative learning	4.60	1.00
D4. Curricular linkage	Weekly topic mapping between curriculum and extracurricular	4.90	1.00
	Post-midterm advanced project linkage	4.10	0.80
D5. AI practice integration	Use of AI laboratory facilities	4.40	1.00
	Cloud-based AI development environment	4.70	1.00
	Generative AI tool utilization	4.90	1.00
	Public data and real-dataset analytics practice	4.40	1.00
D6. Assessment system	Portfolio-based assessment	4.30	1.00
	Output presentation-based assessment	4.60	1.00
	Digital badge and micro-credential linkage	4.10	0.80
D7. Support system	Dedicated coordinator placement	4.40	1.00
	Industry mentor linkage	4.30	1.00
	Participant incentive system	4.60	1.00
	AI equipment and software license provisioning	4.50	1.00

Several patterns merit attention. First, the AI practice integration domain achieved full consensus ($CVR = 1.00$) on all four components, including generative AI tool utilization ($M = 4.90$), confirming that experts view AI-tool engagement as a non-negotiable design element. Second, weekly topic mapping between curriculum and extracurricular activities ($M = 4.90$) and AI tool practical application competency ($M = 4.90$) tied for the highest mean

scores, indicating that experts view curriculum–extracurricular alignment and applied competency as the twin pillars of program design. Third, all four support-system components achieved full consensus, suggesting that experts view operational infrastructure (coordinators, mentors, incentives, equipment) as foundational to program sustainability.

Nine items failed to meet consensus thresholds and were eliminated. Notably, “AI ethics/citizenship debate” was eliminated as a stand-alone program type (CVR = 0.60) but retained as a learning objective (CVR = 0.80), suggesting that experts view AI ethics as a transversal content area rather than a discrete program format. The item “course credit bonus for extracurricular participation” received a CVR of -1.00 (i.e., universal rejection), indicating strong expert consensus that the voluntary character of extracurricular participation should not be compromised by formal credit incentives.

4.3. AHP Results: Direct Quantification of Delphi-Consensus Domains

4.3.1. Weights of Level 1 Evaluation Criteria

The six Level 1 criteria—directly mapped from the six non-program-type Delphi domains—were subjected to pairwise comparison. The aggregated comparison matrix (geometric mean of 10 expert judgments) yielded the weights and consistency metrics presented in Table 2. The consistency ratio of 0.015 confirms acceptable logical consistency well within Saaty’s [23] threshold of 0.10.

Table 2. Weights of Level 1 evaluation criteria (N = 10).

Code	Criterion (Delphi Domain Mapped)	Weight	Rank	Share (%)
C4	AI Practice Integration (D5)	0.286	1	28.6%
C3	Curricular Linkage (D4)	0.286	1 (tie)	28.6%
C1	Learning Objectives (D2)	0.169	3	16.9%
C2	Mode of Operation (D3)	0.112	4	11.2%
C5	Assessment System (D6)	0.088	5	8.8%
C6	Support System (D7)	0.059	6	5.9%

Note: CR = 0.015, $\lambda_{\max}$ = 6.094, CI = 0.019. All criteria correspond directly to Delphi-consensus domains.

AI practice integration (C4 = 0.286) and curricular linkage (C3 = 0.286) emerged jointly as the two highest-weighted criteria, sharing the top rank with identical weights, and jointly accounting for 57.2% of the total weight. This finding empirically confirms the two foundational design principles for sustainable AI/data extracurricular programs in higher education: (i) authentic integration of AI practice (including cloud-based and generative AI tools) and (ii) systematic linkage with formal coursework. Notably, all six criteria correspond directly to Delphi-consensus domains, ensuring that the quantitative prioritization is a direct extension of qualitative consensus rather than the introduction of externally derived evaluative criteria.

Learning objectives (C1 = 0.169) emerged as the third-most-weighted criterion, occupying a position between the dominant practice/linkage criteria and the more infrastructure-oriented criteria. Mode of operation (C2 = 0.112), assessment system (C5 = 0.088), and support system (C6 = 0.059) received progressively lower weights, suggesting that operational and infrastructural considerations—while consensually endorsed as essential components (Table 1)—are viewed by experts as secondary to the pedagogical core of practice integration, curriculum alignment, and learning objectives.

4.3.2. Weights of Level 2 Program Type Alternatives

Pairwise comparison of the three program types under each Level 1 criterion yielded the criterion-specific weights presented in Table 3. All comparison matrices satisfied the consistency threshold ($CR \leq 0.1$). Multiplication of Level 1 weights by Level 2 weights produced the global weights and final priority rankings.

Table 3. Local and global weights of program type alternatives.

Program Type	C1 (0.169)	C2 (0.112)	C3 (0.286)	C4 (0.286)	C5 (0.088)	C6 (0.059)	Global
A1. AI tool workshops	0.297	0.500	0.600	0.297	0.250	0.500	0.414
A2. AI-based problem-solving challenges	0.540	0.250	0.200	0.540	0.500	0.250	0.390
A3. Data analytics projects	0.163	0.250	0.200	0.163	0.250	0.250	0.196

Note: Local weights for each criterion sum to 1.0. Global weight = Sum (Level 1 weight $\times$ Level 2 local weight). All comparison matrices $CR \leq 0.10$.

AI tool workshops (A1) emerged as the top-ranked program type (global weight 0.414), driven primarily by strong performance on curricular linkage and mode of operation—reflecting the workshop format’s natural compatibility with weekly curriculum mapping and team-based blended operation. AI-based problem-solving challenges (A2) ranked a close second (0.390), excelling in AI practice integration, learning objectives, and assessment—reflecting the challenge format’s capacity to integrate deeper AI engagement and deliverable-based assessment. Data analytics projects (A3) ranked third (0.196), positioned as a supplementary type focused on data-centric depth; the underlying local weights are reported in Table 3.

The narrow margin between A1 and A2 ($\Delta = 0.025$) is itself a substantive finding: rather than indicating a single dominant program type, the AHP results suggest a complementary architecture in which workshops serve as the curriculum-aligned core, and challenges serve as the advanced practice-integrated extension. Data analytics projects function as a specialized supplementary track. Given the small margin between the two highest-ranked alternatives, however, this ordering should be read as indicating near-equivalent priority between workshops and challenges rather than a definitive hierarchy, and the “three-tier” characterization is proposed as an interpretive design heuristic rather than an empirically established ordinal structure. So, when interpreted, the structure resonates with UNESCO’s [4] Understand-Apply-Create progression and with Kuh’s [14] tiered high-impact practice typology.

4.4. Content Validity Ratio (CVR) Verification

Following the development of the integrated framework (24 components + AHP priority architecture), CVR verification was conducted to validate the content validity of the consolidated model. Of the 30 verification items organized into 10 domains, 22 items (73.3%) achieved $CVR \geq 0.62$ in the initial verification, with an initial CVI of 0.71. Eight items required revision based on expert qualitative feedback, addressing concerns regarding self-directed learning depth, team composition mechanisms, evaluation balance, audience-vote framing, and crisis-management coverage. For example, the self-directed learning item was reworded to specify guided scaffolding levels; the team composition item was revised to clarify the team assignment mechanism; and the audience-vote item was reframed as a supplementary rather than a primary evaluation element. After these revisions, all

30 items met the CVR threshold in a re-verification round (CVI = 0.84, exceeding Davis's [38] threshold of 0.80). The final framework was thereby confirmed as content-valid.

Table 4 summarizes the verification outcomes by domain. The Program Structure, Contest Topics, and Standard Forms domains achieved the highest average CVRs (0.93 each), confirming strong expert consensus on the fundamental design and standardization of the framework. Domains with lower initial CVRs—including Learning Content (0.67), Final Presentation (0.70), and Performance Management (0.73)—were strengthened through targeted revisions following expert feedback.

Table 4. CVR verification results by domain (N = 10).

Domain	Items	Accepted (Initial)	Mean CVR	Final Judgment
Program structure	3	3	0.93	Strong
Learning content	3	1	0.67	Adequate
Contest topics	3	3	0.93	Strong
Participation and registration	3	2	0.80	Strong
Evaluation system	4	3	0.85	Strong
Final presentation	2	1	0.70	Adequate
Incentives	3	2	0.80	Strong
Standard forms	3	3	0.93	Strong
Operational manual	3	2	0.80	Strong
Performance management	3	2	0.73	Adequate

5. Discussion

5.1. Methodological Contribution: Direct Continuity Between Delphi and AHP

A central methodological contribution of this study is the demonstration of complete continuity between the Delphi and AHP phases, achieved by directly mapping Delphi-consensus domains to AHP hierarchy levels. In many Delphi–AHP studies in educational research, the AHP Level 1 criteria are derived from external sources (e.g., literature-based theoretical frameworks or policy documents), with the result that the qualitative consensus-building phase and the quantitative prioritization phase function as relatively independent analytical steps. The present study instead structured the AHP hierarchy to inherit directly from the Delphi consensus: six of the seven Delphi-consensus domains served as Level 1 evaluation criteria, while the seventh (program type) served as Level 2 alternatives.

This methodological choice carries two important benefits. First, it eliminates the risk of introducing externally derived or post-hoc criteria that could bias the quantitative prioritization. Second, it allows the AHP analysis to function as a direct quantification of the Delphi consensus rather than as an independent analytical step, strengthening the internal coherence of the mixed-methods design. A corollary of this design, however, is that the AHP results should be understood as a quantitative elaboration of the Delphi consensus rather than as an independent validation of it; the epistemic status of the resulting priorities, and the associated risk of consensus self-confirmation, are examined critically in the Limitations. The resulting six-criterion structure remained within Saaty's [23] recommended range (7 ± 2), allowing cognitively tractable pairwise comparison while preserving full domain coverage. This approach may offer a useful reference point for future Delphi–AHP applications in educational program design, although the claimed methodological

advantage should be corroborated through explicit comparison with designs employing externally derived criteria.

5.2. Empirical Validation of Core Design Principles

The AHP results provide robust empirical support for two foundational design principles of sustainable AI/data extracurricular programs in higher education: (i) authentic AI practice integration and (ii) systematic curriculum–extracurricular linkage. These two criteria emerged jointly at the top of the priority ranking, each receiving a weight of 0.286 and together accounting for 57.2% of the total criterion weight. This dual prioritization is theoretically substantive: practice integration without curricular linkage risks producing fragmented experiential learning disconnected from formal coursework, while curricular linkage without practice integration risks reducing extracurricular programs to mere extensions of classroom instruction. The combined emphasis on both criteria reflects an expert vision of extracurricular programs as integrative bridges that connect classroom learning with authentic AI engagement.

This finding extends prior scholarship in two ways. First, it provides quantitative validation of theoretical claims from Kolb’s experiential learning theory [11] and Kuh’s high-impact educational practices framework [14], which posit the necessity of integrated curricular-extracurricular learning. Second, it specifies practice integration and linkage as the two highest-weighted design criteria for AI/data education specifically—a domain where authentic engagement with rapidly evolving tools is particularly critical.

The intermediate weight assigned to learning objectives ($C1 = 0.169$) and the lower weights assigned to mode of operation ($C2 = 0.112$), assessment system ($C5 = 0.088$), and support system ($C6 = 0.059$) reflect expert recognition that operational and infrastructural considerations, while indispensable, are secondary to the pedagogical core. Importantly, this hierarchy does not imply that lower-weighted criteria are dispensable; the Delphi consensus confirmed all four support-system components with $CVR = 1.00$, indicating that experts view these components as essential foundations. Rather, the AHP weights reflect relative emphasis when allocating attention and resources across criteria.

5.3. Complementary Program Type Architecture

The narrow margin between the top two program types in the AHP results (AI tool workshops at 0.414 and AI-based problem-solving challenges at 0.390, $\Delta = 0.025$) suggests that these formats should be conceived as complementary rather than competing. The two formats exhibit distinctive strengths under different criteria: AI tool workshops excel in curricular linkage (local weight 0.600) and mode of operation (0.500), reflecting the workshop format’s natural compatibility with weekly curriculum mapping and regular semester-based operation. AI-based problem-solving challenges, by contrast, excel in AI practice integration (0.540), learning objectives (0.540), and assessment system (0.500), reflecting the challenge format’s capacity to integrate deeper AI engagement, comprehensive learning outcomes, and tangible deliverable-based assessment.

This complementarity supports a tiered program architecture: workshops as the curriculum-aligned core, challenges as the advanced practice-integrated extension, and data analytics projects as a specialized supplementary track. Such an architecture aligns with UNESCO’s [4] developmental Understand–Apply–Create framework and with Kuh’s [14] high-impact practices typology. The key contribution of the present study is the empirical demonstration that these conceptual tiers can be operationalized through three specific program types, with quantified priority rankings to guide institutional resource allocation.

5.4. Sustainable AI Literacy and SDG 4

The findings of this study contribute to the operationalization of SDG 4 (Quality Education) [5] in the context of higher education's response to the AI transition. SDG 4 calls for inclusive, equitable, and lifelong learning opportunities, and the integration of AI literacy into university education is a direct contribution to this goal. However, the sustainability of AI literacy education depends on more than the existence of AI courses; it requires the development of multi-modal learning environments that combine theoretical instruction with experiential engagement, ethical reflection, and authentic problem-solving.

The 24 components identified in this study provide a concrete operationalization of sustainable AI literacy education. This contribution can be made explicit by mapping framework components onto specific SDG 4 targets. Target 4.3 (equal access to affordable, quality tertiary education) is operationalized through the voluntary, open participation structure and through cloud-based practice environments that lower institutional entry barriers; Target 4.4 (relevant skills for employment, decent jobs, and entrepreneurship) is addressed by the AI tool practical application and data interpretation objectives, discipline-specific AI integration, industry mentor linkage, and digital badge and micro-credential recognition; and Target 4.7 (knowledge and skills needed to promote sustainable development, including responsible citizenship) is operationalized through the AI ethical judgment objective and public-data-based analytics practice oriented toward civic problem contexts. The framework's SDG 4 alignment is thus realized through identifiable components rather than asserted at the level of general intent. Several components carry particular significance for sustainability. First, the prioritization of generative AI tools and cloud-based AI development environments (both components of the highest-weighted Level 1 criterion, C4) offers institutions a pathway to AI practice integration without the prohibitive capital costs of physical AI infrastructure—a particularly important consideration for under-resourced institutions. This finding has direct relevance to the equity dimension of SDG 4: by indicating that essential AI practice integration may be achievable through accessible cloud-based and generative tools, the framework offers a potential pathway toward narrowing the digital infrastructure gap between well-resourced and resource-constrained institutions.

Second, the consistent endorsement of voluntary participation (and the rejection of credit-based incentives) preserves intrinsic motivation as the foundation of engagement, in line with self-determination theory [28,29]. At the same time, the universal rejection of the credit-bonus item (CVR = −1.00) should be interpreted with attention to context: in the Korean system, extracurricular programs are institutionally defined as non-credit activities, whereas in several other higher education systems credit-bearing co-curricular provision is regarded as good practice. The unanimity observed here may therefore partly reflect system-specific institutional conventions rather than a universal design principle, and an international panel might reach a different conclusion on this component. Third, the linkage with digital badges and micro-credentials connects extracurricular learning to broader lifelong-learning ecosystems, supporting the SDG 4 emphasis on continuous learning opportunities. Fourth, the panel's strong endorsement of industry mentor linkage and public-data-based analytics practice positions the proposed framework within a wider ecosystem of stakeholders—universities, industry, government, and the broader civic community—thereby supporting the partnerships-for-the-goals dimension of the sustainability agenda. Finally, the centrality of generative AI within the highest-weighted criterion carries governance obligations that extend beyond the AI ethical judgment learning objective. Programs implementing the framework should institute explicit responsible-use provisions covering data privacy (precluding the submission of personal or proprietary data to external AI services), academic integrity (disclosure norms for AI-assisted work and clear boundaries between assistance and substitution), and algorithmic bias (critical

appraisal of tool outputs and bias-aware tool selection). Embedding such provisions at the program-policy level operationalizes the ethical dimension of AI literacy in daily practice and mitigates the risks that accompany intensive generative AI use.

5.5. Theoretical and Practical Implications

Theoretically, this study contributes to three intersecting literatures. First, in the AI literacy education literature [1,9,10,25], it extends conceptual frameworks of AI literacy by specifying the structural components through which AI literacy can be cultivated beyond classroom instruction. Whereas prior work has focused on defining what AI literacy is or measuring its dimensions, this study contributes a structural framework for how AI literacy can be developed through curriculum–extracurricular integration. Second, in the extracurricular and co-curricular learning literature [14–17,27], the study extends existing frameworks by applying them to a specific emerging competency domain (AI/data) and by quantifying the relative weights of design criteria. Third, in the Delphi–AHP methodology literature [22–24,33], the study adds to the growing body of applications demonstrating the methodology’s value for educational program design, with particular emphasis on the methodological benefits of direct continuity between qualitative and quantitative phases.

Practically, the 24-component framework offers a concrete checklist for AI/data extracurricular program design and self-assessment. Institutions seeking to develop such programs can use the framework to identify essential components, prioritize implementation, and benchmark existing offerings. The three-program-type priority architecture (workshops as core, challenges as advanced, analytics projects as specialized) provides a phased-implementation roadmap suitable for institutions with varying levels of resources. For institutions unable to implement all 24 components simultaneously, the AHP weights suggest a three-stage sequence. Stage 1 establishes the two highest-weighted criteria—AI practice integration and curricular linkage (jointly 57.2% of criterion weight)—through AI tool workshops built on generative AI and cloud-based environments and mapped to existing course syllabi, the configuration requiring the least dedicated infrastructure. Stage 2 introduces AI-based problem-solving challenges together with the associated deliverable-based assessment components, deepening practice integration. Stage 3 adds data analytics projects and the remaining support-system components (industry mentor linkage, digital badges and micro-credentials), completing the framework as resources permit. For institutions operating in resource-constrained contexts (e.g., non-metropolitan universities, regional campuses), the framework suggests that essential AI practice integration may be achievable through cloud-based environments and generative AI tools without prohibitive infrastructure investments; because this potential was inferred from expert judgment rather than demonstrated in situ, however, it requires empirical verification in genuinely resource-constrained settings before strong equity claims can be made.

6. Conclusions

This study identified and prioritized the key components of AI/data curriculum-linked extracurricular programs in higher education through a sequential Delphi–AHP design with content validity verification. A methodologically distinctive feature was the direct continuity between the Delphi and AHP phases, with the AHP hierarchy inherited entirely from the Delphi consensus. The findings yielded a 24-component framework organized across seven domains, with AI practice integration and curricular linkage emerging jointly as the foundational design priorities (collectively 57.2% of evaluation weight). Three program types—AI tool workshops, AI-based problem-solving challenges, and data analytics projects—were positioned in a complementary tiered architecture. The final framework achieved a CVI of 0.84, confirming content validity.

These findings carry implications for the sustainable advancement of AI literacy in higher education in alignment with SDG 4. By specifying the structural components and design priorities for AI/data extracurricular programs, the study provides a framework with potential transferability for universities seeking to integrate AI literacy education across curricular and extracurricular contexts, although its applicability beyond the Korean context requires empirical confirmation. The framework's emphasis on accessible AI practice (generative AI, cloud environments) and its compatibility with diverse institutional resource levels further supports equitable AI education aligned with sustainability principles.

Several limitations should be acknowledged. First, the expert panel size ($N = 10$), while consistent with established Delphi methodology in educational program development [24,33], remains modest, and broader replication with larger and internationally diverse panels would strengthen the generalizability of the findings. Second, the same expert panel participated in all three phases of the study—Delphi consensus-building, AHP prioritization, and CVR verification. While this design choice preserved judgmental consistency across phases and reflects common practice in Delphi–AHP research [24,33], it introduces a potential circularity in which the panel that generated the framework also evaluated its validity. Several procedural safeguards mitigated this risk: anonymity was maintained throughout to reduce social conformity pressure, the phases were temporally separated so that each round required independent re-engagement with the items, and the CVR phase employed a distinct judgment task (essentiality rating) rather than a repetition of prior ratings. Even with these safeguards, the design cannot rule out the possibility that the three phases jointly reproduced, rather than independently tested, the panel's initial shared assumptions; accordingly, the reported CVI of 0.84 should be read as evidence of internal consensus coherence rather than of external validity, and confirmation of the framework by an independent expert panel remains an essential task for future research. Third, the framework was developed within the Korean higher education system. Korea represents an informative leading-edge case—combining strong national AI education policy with acute demographic and resource pressures on regional universities—but institutional governance, funding structures, and extracurricular traditions differ across countries, and cross-national adaptation of the framework therefore requires further investigation. In particular, the panel's resource-constrained perspective derives from non-metropolitan institutions within a developed higher education system; institutions in low-income countries or without basic AI infrastructure were not represented, and claims regarding resource-constrained applicability are accordingly scoped to relatively under-resourced institutions within developed systems. Fourth, the present study identifies and validates structural components but does not empirically demonstrate the effectiveness of programs designed according to the framework.

These limitations point to important directions for future research. First, empirical implementation studies should evaluate the effectiveness of programs designed using this framework, employing quasi-experimental or longitudinal designs to assess AI competency gains. Second, cross-national replication and adaptation studies should examine the framework's transferability to other higher education contexts. Third, qualitative studies of student experience in framework-based programs would deepen understanding of the lived dimensions of AI/data extracurricular learning. Fourth, longitudinal studies tracking program graduates' AI-related career trajectories would help assess the long-term sustainability impact of curriculum-linked AI literacy education. Finally, once program-level outcome data (e.g., AI competency gains, satisfaction, career outcomes) become available, data-driven feature-selection methods—such as the recently proposed CR-SCAD algorithm, which combines collaborative representation with the smoothly clipped absolute deviation

penalty to identify discriminative indicators while suppressing redundant ones [40]—could complement the expert-derived priorities by empirically testing which of the 24 components most strongly predict program effectiveness, providing a data-driven counterpart to the consensus-based validation employed here. This study thereby lays a foundation for these subsequent inquiries into sustainable AI literacy in higher education.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/su18168050/s1>. Table S1. Delphi Round 2 items eliminated for failing the consensus thresholds (N = 10). Retention required all three criteria: content validity ratio (CVR) ≥ 0.62 , coefficient of variation (CV) ≤ 0.5 , and degree of consensus ≥ 0.75 . Of the 33 candidate items evaluated in Round 2, 24 met all three thresholds and were retained; the nine items below failed at least one threshold and were removed.

Author Contributions: Conceptualization, J.-S.B.; methodology, J.-S.B. and K.-M.K.; formal analysis, J.-S.B.; investigation, J.-S.B. and H.-J.P.; data curation, J.-S.B.; writing—original draft preparation, J.-S.B.; writing—review and editing, K.-M.K. and H.-J.P.; supervision, J.-S.B.; project administration, J.-S.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This study falls within the categories of research exempt from mandatory Institutional Review Board review under the Republic of Korea’s Bioethics and Safety Act and its Enforcement Rule. The research was non-interventional, did not involve human biological materials, tissues, or clinical interventions, and consisted solely of an anonymous, minimal-risk expert survey (Delphi and AHP) completed by voluntary adult professionals. No sensitive personal information was collected, and all responses were collected anonymously. This study therefore falls within the exempt categories as defined by the relevant legislation.

Informed Consent Statement: Informed consent was obtained from all participants involved in the study. Prior to participation, all expert panel members were informed of the purpose of the study, the voluntary nature of participation and the right to withdraw at any time without disadvantage, the anonymous handling of their responses, and the intended academic use of the aggregated results; consent was obtained through their voluntary completion and submission of the questionnaires.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy restrictions concerning the expert panel participants.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Long, D.; Magerko, B. What is AI literacy? Competencies and design considerations. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; pp. 1–16. [CrossRef]
- Lee, Y.M.; Park, Y.S. Establishing a definition of AI literacy and designing a liberal arts education program. *J. Lang. Lit.* **2021**, *85*, 451–474. [CrossRef]
- Chee, H.; Ahn, S.; Lee, J. A competency framework for AI literacy: Variations by different learner groups and an implied learning pathway. *Br. J. Educ. Technol.* **2025**, *56*, 2146–2182. [CrossRef]
- UNESCO. *AI Competency Framework for Teachers*; UNESCO: Paris, France, 2024.
- United Nations. *Transforming Our World: The 2030 Agenda for Sustainable Development*; United Nations: New York, NY, USA, 2015.
- Cardona, M.; Rodriguez, R.J.; Ishmael, K. *Artificial Intelligence and the Future of Teaching and Learning*; U.S. Department of Education: Washington, DC, USA, 2023.
- Knox, J. The political economy of education and the post-digital condition. *Postdigit. Sci. Educ.* **2023**, *5*, 1–17.
- World Economic Forum. *The Future of Jobs Report 2025*; World Economic Forum: Geneva, Switzerland, 2025.
- Ng, D.T.K.; Leung, J.K.L.; Chu, S.K.W.; Qiao, M.S. Conceptualizing AI literacy: An exploratory review. *Comput. Educ. Artif. Intell.* **2021**, *2*, 100041. [CrossRef]
- Chiu, T.K.; Ahmad, Z.; Ismailov, M.; Sanusi, I.T. What are artificial intelligence literacy and competency? A comprehensive framework to support them. *Comput. Educ. Open* **2024**, *6*, 100171. [CrossRef]

11. Kolb, D.A. *Experiential Learning: Experience as the Source of Learning and Development*, 2nd ed.; FT Press: Upper Saddle River, NJ, USA, 2014.
12. Kim, J.M.; Yoo, Y.E.; Lee, S.W. The analysis of effectiveness and review of university extra-curricular programs. *J. Learn.-Cent. Curric. Instr.* **2017**, *17*, 773–791. [[CrossRef](#)]
13. Pascarella, E.T.; Terenzini, P.T. *How College Affects Students: A Third Decade of Research*; Jossey-Bass: San Francisco, CA, USA, 2005; Volume 2.
14. Kuh, G.D. High-impact educational practices. *Peer Rev.* **2008**, *10*, 30–31.
15. Stirling, A.E.; Kerr, G.A. Creating meaningful co-curricular experiences in higher education. *J. Educ. Soc. Policy* **2015**, *2*, 1–7. [[CrossRef](#)]
16. Choi, J.H. The effect of co-curriculum on students' major competency and satisfaction. *J. Educ. Cult.* **2020**, *26*, 515–533. [[CrossRef](#)]
17. Yoon, S.S.; Kim, B.K. Effects of curriculum-linked project-based extracurricular education based on industry-university cooperation: Focusing on D University. *J. Educ. Innov.* **2020**, *30*, 149–174. [[CrossRef](#)]
18. Laupichler, M.C.; Aster, A.; Schirch, J.; Raupach, T. Artificial intelligence literacy in higher and adult education: A scoping literature review. *Comput. Educ. Artif. Intell.* **2022**, *3*, 100101. [[CrossRef](#)]
19. Almatrafi, O.; Johri, A.; Lee, H. A systematic review of AI literacy conceptualization, constructs, and implementation and assessment efforts (2019–2023). *Comput. Educ. Open* **2024**, *6*, 100173. [[CrossRef](#)]
20. Kassorla, M.; Georgieva, M.; Papini, A. *AI Literacy in Teaching and Learning: A Durable Framework for Higher Education*; EDUCAUSE: Louisville, CO, USA, 2024.
21. European Commission. *Empowering Learners for the Age of AI: An AI Literacy Framework for Primary and Secondary Education*; OECD-European Commission: Paris, France, 2025.
22. Green, R.A. The Delphi technique in educational research. *SAGE Open* **2014**, *4*, 2158244014529773. [[CrossRef](#)]
23. Saaty, R.W. The analytic hierarchy process—What it is and how it is used. *Math. Model.* **1987**, *9*, 161–176. [[CrossRef](#)]
24. Tang, H.W.V. Optimizing an immersion ESL curriculum using analytic hierarchy process. *Eval. Program Plan.* **2011**, *34*, 343–352. [[CrossRef](#)] [[PubMed](#)]
25. Ng, D.T.K.; Wu, W.; Leung, J.K.L.; Chiu, T.K.F.; Chu, S.K.W. Design and validation of the AI literacy questionnaire: The affective, behavioural, cognitive and ethical approach. *Br. J. Educ. Technol.* **2024**, *55*, 1082–1104. [[CrossRef](#)]
26. Holmes, W.; Tuomi, I. State of the art and practice in AI in education. *Eur. J. Educ.* **2022**, *57*, 542–570. [[CrossRef](#)]
27. Bartkus, K.R.; Nemelka, B.; Nemelka, M.; Gardner, P. Clarifying the meaning of extracurricular activity: A literature review of definitions. *Am. J. Bus. Educ.* **2012**, *5*, 693–704. [[CrossRef](#)]
28. Deci, E.L.; Ryan, R.M. *Intrinsic Motivation and Self-Determination in Human Behavior*; Springer Science & Business Media: New York, NY, USA, 2013.
29. Ryan, R.M.; Deci, E.L. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* **2000**, *55*, 68–78. [[CrossRef](#)] [[PubMed](#)]
30. Chiu, T.K. The impact of Generative AI (GenAI) on practices, policies and research direction in education: A case of ChatGPT and Midjourney. *Interact. Learn. Environ.* **2024**, *32*, 6187–6203. [[CrossRef](#)]
31. Linstone, H.A.; Turoff, M. *The Delphi Method: Techniques and Applications*; Addison-Wesley: Reading, MA, USA, 1975.
32. Reeves, G.; Jauch, L.R. Curriculum development through Delphi. *Res. High. Educ.* **1978**, *8*, 157–168. [[CrossRef](#)]
33. Yao, Y.; Cao, J.; Tan, Z.; Chen, K. Developing a digital health competency assessment framework for public health services: A Delphi-AHP approach. *Front. Public Health* **2025**, *13*, 1634261. [[CrossRef](#)] [[PubMed](#)]
34. Creswell, J.W.; Plano Clark, V.L. *Designing and Conducting Mixed Methods Research*, 3rd ed.; SAGE Publications: Thousand Oaks, CA, USA, 2017.
35. Lawshe, C.H. A quantitative approach to content validity. *Pers. Psychol.* **1975**, *28*, 563–575. [[CrossRef](#)]
36. Forman, E.; Peniwati, K. Aggregating individual judgments and priorities with the analytic hierarchy process. *Eur. J. Oper. Res.* **1998**, *108*, 165–169. [[CrossRef](#)]
37. Wilson, F.R.; Pan, W.; Schumsky, D.A. Recalculation of the critical values for Lawshe's content validity ratio. *Meas. Eval. Couns. Dev.* **2012**, *45*, 197–210. [[CrossRef](#)]
38. Davis, L.L. Instrument review: Getting the most from a panel of experts. *Appl. Nurs. Res.* **1992**, *5*, 194–197. [[CrossRef](#)]
39. Mesinger, D.; Ociecek, A.; Owczarek, T. Attitudes of young Tri-City residents toward game meat: Development and validation of a scale for identifying attitudes toward wild meat. *Int. J. Environ. Res. Public Health* **2023**, *20*, 1247. [[CrossRef](#)] [[PubMed](#)]
40. Cai, J.; Liu, S.; Wang, B.; Chen, T.; Zhu, N.; Chen, R. Developing a CR-SCAD algorithm for fibrosis and inflammatory activity analysis of chronic hepatitis C. *Int. J. Mach. Learn. Cybern.* **2026**, *17*, 3. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.