Journal of Theoretical and Applied Information Technology
15th May 2026. Vol.104. No.9

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

IMPROVING THE HIGHER EDUCATION ENVIRONMENT THROUGH DETERMINISTIC ASSESSMENT: A CLOUD-INTEGRATED DECISION SUPPORT SYSTEM FOR ACADEMIC ACCREDITATION

MOHAMMED HUSSAIN ALHARBI[1], HAFEDH MAHMOUD ZAYANI[2]

[1]Department of Curricula and Instructional Technology, College of Humanities and Social Sciences, Northern Border University, Arar, Saudi Arabia
[1]ORCID: https://orcid.org/0000-0003-4968-2594
[2]Department of Electrical Engineering, College of Engineering, Northern Border University, Arar, Saudi Arabia
[2]ORCID: https://orcid.org/0000-0001-7195-9743
*Corresponding Author E-mail: [1]Mohammed.alharbi2@nbu.edu.sa, [2]hafedh.zayani@nbu.edu.sa

ABSTRACT

Traditional academic accreditation processes often suffer from subjective bias and non-reproducible outcomes due to a heavy reliance on holistic human judgment, which creates a critical data-integrity gap in quality assurance. This study introduces the Accreditation Standards Evaluation Platform (ASEP), a cloud-integrated Decision Support System (DSS) designed to resolve this problem by transitioning institutional evaluation into a deterministic, rule-based computational framework. At the core of the ASEP architecture is the decomposition of accreditation standards into seven measurable performance dimensions (E1 through E7), which are processed via a unique 'Sum-and-Compare' algorithmic logic. By utilizing a Binary Evaluation Vector (V_{in}) and a sequential matrix validation process—incorporating a dedicated consistency layer (E6 to cross-validate application and evidence)—the system eliminates the qualitative "black-box" nature of traditional assessments. The pilot implementation results demonstrate that the ASEP system achieved a 60% reduction in assessment processing time compared to manual methods. More importantly, the system's deterministic logic resulted in a 66% decrease in missing evidence and a 30% reduction in inflated compliance ratings, effectively bridging the 62% average discrepancy previously observed between self-assessment and independent audit scores. This research confirms that algorithmic determinism is a necessary evolution for achieving scalable, audit-ready quality assurance in higher education.

Keywords: *Decision Support Systems; Rule-Based Algorithms; Automated Assessment; Quality Assurance; Educational Data Mining; System Architecture*

1. INTRODUCTION

In the domain of complex institutional evaluation, the reliability of Decision Support Systems (DSS) is fundamentally contingent upon the implementation of rigorous, consistent, and transparent computational mechanisms. While multi-criteria decision-making (MCDM) and AI-assisted models have advanced the automation of assessment processes, they often rely on probabilistic inference or expert-defined weights that introduce variability and interpretive discretion. In high-stakes environments such as academic accreditation, where data integrity directly influences strategic planning and institutional benchmarking, the lack of deterministic logic remains a critical systemic vulnerability. Traditional evaluation frameworks frequently suffer from the "black-box" nature of holistic judgments, where heterogeneous performance indicators are prematurely aggregated into single compliance scores, obscuring specific institutional gaps and undermining reproducibility.

In the Kingdom of Saudi Arabia, the academic accreditation process is governed by the Education and Training Evaluation Commission (ETEC) through the National Centre for Academic Accreditation and Evaluation (NCAAA). Although the NCAAA employs a sophisticated hierarchical quality model comprising standards, criteria, and elements, the practical application of this model faces significant computational challenges. The existing self-evaluation process relies heavily on subjective, criterion-level judgments, leading to inconsistent outcomes that compromise the reliability of data-driven decision-making. Current evaluation

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

methodologies lack fixed logical rules aligned with compliance requirements. Consequently, evaluators frequently merge distinct factors—such as policy existence, implementation quality, and evidence availability—into a single rating through non-deterministic aggregation, which obscures specific institutional gaps. This approach frequently results in "false positives," where critical deficiencies in implementation consistency are masked by strengths in secondary areas, leaving program leaders without the precise diagnostic intelligence required for effective resource allocation.

The significance of addressing this methodological vulnerability extends far beyond regional accreditation practices. Globally, academic quality assurance operates as a high-stakes regulatory mechanism where evaluation outcomes directly influence institutional funding, national rankings, and public trust. When assessment models rely on subjective aggregation or probabilistic inference, they introduce systemic variability that compromises data integrity and obscures actionable institutional gaps. This 'Logic Layer' vulnerability allows peripheral strengths—such as well-written policies—to statistically mask critical compliance failures in implementation or evidence, creating 'false positives' that misdirect strategic resources. In an era where educational data analytics drive institutional decision-making, the absence of deterministic, auditable evaluation logic represents a critical vulnerability. Transitioning to rule-based computational architectures is therefore not merely a technical optimization, but a necessary evolution to ensure transparency, reproducibility, and evidence-based institutional improvement at scale.

To address these systemic gaps, this research proposes a deterministic, rule-based diagnostic framework designed to transform self-evaluation from a static compliance exercise into a dynamic, actionable computational tool. Unlike traditional systems that rely on fuzzy logic or weighted averages, this framework enforces a 'Non-Negotiable Quality Floor' via an Essential Criteria Constraint. This ensures that accreditation levels are strictly contingent upon meeting foundational requirements, preventing statistical averaging from compensating for core compliance failures. The framework operationalizes this transition by deconstructing each accreditation criterion into measurable elements and evaluating them across six primary, interconnected dimensions using Binary Evaluation Vectors (V_{in}) and sequential matrix validation.

The core contribution of this work lies in how the framework mathematically formalizes assessment logic. By utilizing six mathematical formulas and two logical decision matrices, the framework replaces interpretive scoring with deterministic, reproducible rules. This logic is fully operationalized within the Accreditation Standards Evaluation Platform (ASEP), a cloud-integrated system architecture that mandates a direct linkage between numerical scores and verifiable evidence. The system enforces a 'Sum-and-Compare' algorithmic logic that eliminates subjective weighting, ensuring full reproducibility without interpretive discretion.

The contributions of this work are significant from both a system design and application perspective:

1. Formalizing Scoring Logic: The system converts qualitative accreditation standards into fixed mathematical rules, eliminating subjective weighting and enhancing data integrity.
2. Diagnostic Granularity: It enhances diagnostic utility by distinguishing between different types of failure (e.g., missing policy vs. poor implementation) through dimensional separation.
3. Empirical Validation: Pilot data demonstrates the framework's efficacy, showing a 62% average reduction in score discrepancies, a 66% decrease in missing evidence, and a 60% reduction in assessment processing time.

By correcting the compliance distribution—specifically reducing inflated ratings by 30% through the elimination of 'false positives'—the system ensures that final outcomes are a true reflection of institutional quality. These results confirm that the transition to deterministic logic effectively bridges the gap between subjective self-assessment and objective audit results, providing the high-fidelity diagnostic intelligence required for institutional improvement.

The remainder of this paper is organized as follows: Section 2 reviews the literature on global quality frameworks and existing computational assessment models, concluding with the identification of current logic gaps. Section 3 explicitly defines the research problem and the systemic vulnerabilities in traditional scoring. Section 4 details the NCAAA structure and the proposed framework enhancements. Section 5 presents the Research Method Protocol, including the quantitative methodology, decision matrices, and the ASEP algorithmic logic. Section 6 describes the system implementation and validation results, while Section 7 provides a comparative critique against state-of-the-art models. Finally, Section 8 offers concluding remarks and outlines future research directions.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2. LITERATURE REVIEW

The computational management of quality assurance in higher education represents a complex Decision Support System (DSS) challenge, where data integrity and algorithmic consistency are paramount. Globally, accreditation frameworks such as ABET [16], the UK's QAA [12], and the European ESG [6] prioritize systematic quality management. A comprehensive literature review by Duarte and Vardasca [4] highlights that while accreditation systems are widespread, their underlying assessment methodologies often lack standardization, leading to variability in outcomes across different jurisdictions. In Saudi Arabia, the Education and Training Evaluation Commission (ETEC) governs this process through the National Centre for Academic Accreditation and Evaluation (NCAAA), which defines a hierarchical model of standards, criteria, and evidence expectations [5]. While these frameworks provide the policy layer for quality assurance, their implementation layer often relies on manual, subjective processes that introduce variability into institutional data.

Recent literature has attempted to reduce subjective evaluations by applying Multi-Criteria Decision Making (MCDM) and AI-assisted models. Alqahtani et al. [1] addressed NCAAA subjectivity by decomposing criteria into sub-criteria and applying TOPSIS for program ranking. Their work demonstrated that finer-grained assessment yields more conservative—and arguably more accurate—ratings. However, while decomposing criteria improves granularity, relying on TOPSIS reintroduces subjectivity through expert-defined weights. Similarly, traditional approaches often utilize Analytic Hierarchy Process (AHP) [13], [15] or Fuzzy Logic [2], which depend on continuous expert input and probabilistic inference to handle uncertainty.

While MCDM and Fuzzy Logic offer flexibility, they present a 'black-box' risk in regulatory contexts where auditability is required. In accreditation, where legal and institutional accountability are paramount, probabilistic models can obscure the reasoning behind specific compliance decisions. This research posits that a deterministic rule-engine offers superior utility for accreditation than probabilistic models. By formalizing the NCAAA rubric into fixed mathematical rules with predefined thresholds, this study ensures fully deterministic, reproducible scoring without human bias. Unlike weighted models where high scores in secondary areas can compensate for core failures, a rule-based approach enforces logical constraints that mirror regulatory compliance requirements.

From a system design perspective, existing quality assurance platforms often function primarily as data repositories rather than intelligent assessment engines. Naim et al. (2024) [10] proposed information system models for learning outcomes, yet these often rely on manual data entry and lack embedded logical validation. Similarly, KPI-based dashboards [7], [9] provide visualization but rarely enforce logic-based constraints at the point of data entry. These systems address the presentation layer (reporting) but neglect the logic layer (computational assessment rules).

The gap lies in how educational platforms integrate Business Rule Management Systems (BRMS). Unlike general workflow systems, the proposed framework embeds the assessment logic directly into the system architecture, preventing invalid states rather than merely reporting them post-hoc. While some studies have explored computational tools like Natural Language Processing (NLP) for evidence detection [11], these focus on efficiency or evidence extraction—not on reforming the core assessment logic. The integrity of institutional data is a foundational challenge in educational analytics; traditional processes suffer from data heterogeneity, where qualitative narratives are inconsistently mapped to quantitative scores. By formalizing the scoring rubric into mathematical vectors, this work contributes to the field of Quality Assurance Analytics, ensuring that data generated during self-evaluation is structured, consistent, and suitable for higher-order computational analysis.

Critically, a technical gap remains in the algorithmic formalization of accreditation logic. Existing literature addresses either the policy layer (standards definition) or the presentation layer (dashboards and reporting), but neglects the logic layer. No existing work provides a comprehensive system that: (1) decomposes criteria into measurable data elements, (2) processes them through a deterministic logical engine without subjective weighting, and (3) enforces structural constraints (e.g., Essential Criteria) at the system architecture level.

This research fills this computational gap by introducing a fully operationalized, evidence-anchored rule engine that transforms self-evaluation from a data entry task into a diagnostic computing process. By shifting the focus from subjective aggregation to algorithmic determinism, this work provides a robust foundation for the next generation of intelligent evaluation systems, ensuring that institutional excellence is built on verifiable data rather than interpretive judgment [3], [8].

While this study utilizes the NCAAA standards as a primary testbed, the challenges of subjective assessment are a global phenomenon documented across diverse regulatory environments. International frameworks such as the Accreditation Board for Engineering and Technology (ABET) in the United States and the Quality Assurance Agency (QAA) in the United Kingdom similarly struggle with 'evaluator variability.' For instance, the UK Quality Code emphasizes evidence-based outcomes, yet often lacks a deterministic computational layer to verify the consistency between reported KPIs and actual implementation. By connecting the ASEP architecture to these global standards, this research addresses a universal requirement for reproducible, high-fidelity quality data that transcends national boundaries.

3. PROBLEM STATEMENT

Despite the proliferation of cloud-based Management Information Systems (MIS) in higher education, institutional assessment remains trapped in a 'Logic Layer' gap where qualitative narratives are inconsistently mapped to quantitative scores. While current systems serve as efficient data repositories, they lack the computational intelligence to programmatically detect discrepancies between 'reported' compliance and 'actual' evidence availability [7], [9], [10].

This problem is evidenced by the prevalence of "false positives" in accreditation outcomes. During the 2022/2023 cycle at our institution, manual audits revealed that 18% of criterion scores contained internal logical contradictions—such as "Level 4" ratings being assigned despite the documented absence of regular assessment. These contradictions are often exacerbated by traditional weighted-averaging models (e.g., AHP, TOPSIS, Fuzzy Logic) which dominate existing literature [1], [2], [13]. These models allow peripheral strengths, such as well-written policies, to statistically mask critical failures in implementation, effectively creating a "black-box" that obscures institutional gaps and misdirects strategic resources [4].

Consequently, there is an urgent need for a deterministic Decision Support System (DSS) capable of: (1) mathematically validating compliance against fixed logical rules, (2) enforcing cross-dimensional consistency constraints to prevent statistical masking, and (3) replacing subjective aggregation with a reproducible, rule-based computational workflow. Without such a framework, accreditation results remain non-reproducible narratives rather than audit-ready diagnostic intelligence.

4. FRAMEWORK DESIGN AND THEORETICAL BASIS

This section outlines the official ETEC-NCAAA accreditation structure [5] and introduces the methodological enhancements that enable a quantitative, diagnostic self-evaluation framework. The model respects NCAAA’s hierarchical design while introducing operational innovations for precise, reproducible assessment.

4.1. Hierarchical Structure of NCAAA Standards

The NCAAA quality model employs a three-tier structure:

- Standard (*Lstandard*): Broad domains (e.g., Teaching and Learning).
- Criterion (*Lcriterion*): Specific, measurable requirements; some are essential (mandatory).
- Evidence Expectations: General documentation guidance—no formal scoring rules are provided, leaving evaluation reliant on subjective interpretation.

4.2. Framework Enhancements

Three innovations transform this qualitative structure into a quantitative system

- Element-Level Decomposition: Criteria are split into discrete, measurable elements—the unit of assessment.
- Six-Dimensional Assessment: Each element is evaluated across: E1 (Elements Availability), E2 (Quality of Application), E3 (Regularity of Application), E4 (Regularity of Assessment), E5 (Evidence of Application and Evaluation Availability), E7 (Continuous Improvement)
- Evidence Guidance System: Practical, accreditation-informed checklists standardise evidence collection without restricting flexibility.

4.3. The Six Assessment Dimensions

The framework operationalizes quality assessment through six distinct dimensions, each measuring specific aspects of program performance:

- E1 (Element Availability): Assesses whether the required components, resources, and policies related to the criteria are present and formally approved. This dimension answers the fundamental question:” Do the necessary elements exist?”
- E2 (Application Quality): Assesses the effectiveness, appropriateness, and

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

implementation quality of practices or policies in fulfilling the criterion's intended purpose.

- E3 (Application Regularity): Measures the consistency, frequency, and sustained application of practices or policies over the required period.
- E4 (Assessment Regularity): Evaluates the consistency and effectiveness of monitoring and evaluation mechanisms for criterion implementation.
- E5 (Evidence of Application and Evaluation Availability): Examines the existence and quality of documentation that proves implementation practices and assessment activities. This dimension answers:" Can we demonstrate how elements are being used and evaluated?"
- E7 (Continuous improvement and level of results in the light of indicators and benchmarking): Assesses whether evidence is systematically collected, results are analysed, and findings are formally used to modify and improve practices or policies.

The clear separation between E1 (Elements Availability) and E5 (Evidence of Application and Evaluation Availability) is particularly crucial, as it distinguishes between the existence of required components and the documentation of their implementation—a distinction often blurred in traditional assessments.

It is important to note that E6 is not an independent input dimension but rather a composite index calculated from E3, E4, and E5 via Matrix 1. This composite serves as a consistency check within the decision logic.

4.4. *Theoretical Foundations of Dimensional Assessment*

Although rule-based, the framework is implicitly grounded in learning theory. Bloom's Taxonomy informs E2 by targeting the "Apply" level through observable implementation evidence. Situated Learning Theory supports the shift from policy awareness to active participation in quality processes. Vygotsky's scaffolding underpins the ASEP platform's role in developing faculty assessment literacy through structured, evidence-based engagement [14]. Together, these theories justify the move from passive compliance to developmental self-evaluation—without compromising the framework's deterministic logic.

4.5. *Foundations for Quantitative Scoring*

The core innovation is the mathematical formalization of NCAAA assessment through six formulas and two decision matrices. The framework utilizes six primary dimensions (E1–E5, E7) and one derived composite dimension (E6) to calculate the final compliance level. Elemental scores are aggregated into criterion and standard level compliance ratings using fixed, reproducible rules—eliminating interpretive judgment while preserving NCAAA's quality philosophy. This formalization enables the diagnostic precision and objectivity.

5. METHODOLOGY

To ensure evaluative objectivity, the framework applies a hierarchical filtering logic that processes data through two sequential decision matrices. First, the Evidence-Application Matrix (Matrix 1) evaluates the consistency between application regularity (E3), assessment regularity (E4), and evidence availability (E5). This ensures that a criterion cannot receive a high compliance rating if practices are applied sporadically, rarely assessed, or poorly documented—even if implementation quality (E2) is high. Second, the Compliance-Impact Matrix (Matrix 2) integrates all six primary dimensions (E1, E2, E3, E4, E5, E7) to determine the final criterion-level compliance. Note that E6 is an intermediate composite used internally but not included as an input dimension in the final compliance vector V_{in}. This dual-layered approach prevents 'false positives' in scoring, where a program might claim excellence without having the documentation or the 'closing the loop' evidence to prove it.

5.1. *From Subjective Aggregation to Multi-Dimensional Assessment*

The framework builds upon, yet significantly enhances, the foundational principles outlined in the Self-Evaluation Scales for Higher Education Institutions document (Education and Training Evaluation Commission, 2021). The original methodology employed a five-point scale (1 to 5) with Level 3 as the compliance threshold, requiring evaluators to subjectively aggregate multiple performance aspects into a single criterion score.

The fundamental limitation of this approach was the loss of diagnostic specificity. Critical weaknesses in specific dimensions (e.g., poor regularity or absent continuous improvement) could be obscured by strengths in other areas, preventing targeted interventions. The proposed framework addresses this through systematic decomposition and mathematical formalization:

- Dimensional Separation: Isolating six distinct assessment dimensions to prevent compensation effects

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

- Element-Level Granularity: Decomposing criteria into measurable elements for precise assessment .
- Deterministic Aggregation: Replacing subjective judgment with mathematical formulas and logical matrices.

5.2. The Six Assessment Dimensions

Each criterion is decomposed into its constituent elements, which are evaluated across six dimensions using predefined qualitative descriptors. Each descriptor is mapped to a quantitative weight (w), ranging from 0 to 4, as detailed in Table 1. This structured scoring system ensures that qualitative observations are converted into a standardized format suitable for rule-based diagnostic analysis.

To illustrate the diagnostic value, consider a hypothetical criterion where all policies exist (E1 = 4), but implementation is inconsistent (E3 = 2) and evidence is weak (E5 = 1). Traditional scoring might average these into a misleading ‘Level 3’. In contrast, the rule-based framework flags this as Level 2 due to failed consistency checks—directing attention to operational gaps rather than policy existence.

Table 1. Dimensional Assessment Criteria and Weighting Logic

Dim	Qualitative Descriptors	*w*
E1	*Item Available*	4
	Item Not Available	0
E2	*Elements applied at a perfect level*	4
	Elements applied at a good level	3
	Elements applied at a low level	2
	Elements not applied, or applied at a very low level	1
E3	*Applied regularly*	4
	Applied irregularly	2
	Rarely applied	1
E4	*Regular and effective assessment*	4
	Regular assessment	3
	Irregular assessment	2
	No assessment	1
E5	*Sufficient and varied evidence available*	4
	Sufficient evidence available	3
	Insufficient evidence available	2
	No evidence available	1
E7	*Regular procedures with measurable improvement*	4
	Regular procedures with good results	3
	Limited improvement procedures	2
	No improvement procedures in place	1

Having established the scoring logic per dimension, the framework then generates a composite dimension—Combined Consistency and Evidence (E6)—to provide a holistic view of performance. This is achieved by synthesizing the categorical interpretations of Application Regularity (E3), Assessment Regularity (E4), and Evidence of Application and Evaluation Availability (E5).

5.3. Quantitative Workflow and Research Protocol

The framework follows a systematic seven-step research protocol process that transforms qualitative evidence into reproducible quantitative assessments through a deterministic logic layer:

- **Criterion Decomposition**: Break down each NCAAA criterion into atomic, measurable elements to eliminate interpretive ambiguity.
- **Element-Level Assessment (Vector Initialization):** Evaluate each element across the fundamental dimensions (E1-E5, E7). This stage serves as the initialization of the **Binary Evaluation Vector (*V_{in}*)** based on predefined qualitative descriptors.
- **Dimensional Scoring:** Calculate percentage achievements and assign qualitative levels using mathematical formulas. This step transitions the data from raw observation to a **deterministic logic state** using dimension-specific thresholds.
- **Composite Integration (Logic Processing):** Apply **Matrix 1** to generate the E6 composite dimension. This represents the core "Consistency Layer," logically cross-referencing regularity and evidence (E3, E4, E5) to prevent "false positive" compliance.
- **Final Criterion Determination:** Apply **Matrix 2** to integrate all dimensional assessments into final compliance levels. This acts as the **final algorithmic decision gate** for each criterion.
- **Standard-Level Aggregation:** Calculate standard scores while programmatically enforcing **Essential Criteria Constraints**. If an essential criterion is not met, the global status is automatically restricted to ensure quality integrity.
- **Data Validation and Cloud Sync:** The final output is synchronized with the cloud-integrated dashboard for multi-tier verification, ensuring the results are transparent and auditable.

5.4. Core Formulas and Matrices

5.4.1. Formula 1: Criterion Achievement Percentage Calculation

For a specific criterion, the achievement percentage for each assessment dimension $\boldsymbol{d} \in \{E1, E2, E3, E4, E5, E7\}$ is calculated by aggregating the scores (w: weight) of its constituent elements as follows:

$$\boldsymbol{P_d} = \frac{\sum_{k=1}^{n} \boldsymbol{w_{k,d}}}{\boldsymbol{4n}} \times \boldsymbol{100\%} \quad (1)$$

Where:

- P_d: The achievement percentage for dimension d relative to the specific criterion being assessed.
- $w_{k,d}$: The weight assigned to the qualitative descriptor for element k within that criterion, relative to dimension d (refer to Table 1).
- n: The total number of elements constituting the given criterion
- $4n$: The maximum possible score for the criterion in that dimension (representing a perfect score of 4 for all n elements).

5.4.2. Formula 2: Dimension Level Mapping and Categorical Assignment

Following the calculation of the achievement percentages, the framework assigns a discrete qualitative level (1 to 4) to each dimension relative to the specific criterion. This step is essential for converting continuous quantitative data (P_d) into the categorical interpretations required for the logical matrix operations in the subsequent stages. The assignment of the qualitative level L_d for a given criterion is defined as follows:

$$L_d = \begin{cases} 4 & T_{d,4} \le P_d \\ 3 & T_{d,3} \le P_d < T_{d,4} \\ 2 & T_{d,2} \le P_d < T_{d,3} \\ 1 & T_{d,1} \le P_d < T_{d,2} \end{cases} \quad (2)$$

Where:

- L_d: The resulting qualitative level (1, 2, 3, or 4) for dimension d relative to the assessed criterion.
- P_d : The achievement percentage for the criterion calculated in Equation 1.
- $T_{d,j}$: The dimension-specific threshold values (defined in Table 2) used to determine the boundary for level j.

Table 2. Criterion-Level Threshold Mapping for Qualitative Level Assignment

Dim	Level 4	Level 3	Level 2	Level 1
E1	$T_{1,4} = 100\%$	$T_{1,3} = NA^*$	$T_{1,2} = 60\%$	$T_{1,1} = 0\%$
E2	$T_{2,4} = 80\%$	$T_{2,3} = 60\%$	$T_{2,2} = 40\%$	$T_{2,1} = 0\%$
E3	$T_{3,4} = 100\%$	$T_{3,3} = NA^*$	$T_{3,2} = 40\%$	$T_{3,1} = 0\%$
E4	$T_{4,4} = 80\%$	$T_{4,3} = 60\%$	$T_{4,2} = 40\%$	$T_{4,1} = 0\%$
E5	$T_{5,4} = 80\%$	$T_{5,3} = 60\%$	$T_{5,2} = 30\%$	$T_{5,1} = 0\%$
E7	$T_{7,4} = 80\%$	$T_{7,3} = 60\%$	$T_{7,2} = 40\%$	$T_{7,1} = 0\%$

**NA: indicates that the corresponding level is not used in the threshold logic.*

5.4.3. Formula 3: Qualitative Interpretation Mapping

The final step in the dimensional analysis is the translation of numerical levels (L_d) back into standardized qualitative narratives for each specific criterion. This mapping ensures that the results are accessible to stakeholders and align with the descriptive requirements of international accreditation reports. By converting the calculated status into a qualitative descriptor, the framework maintains a consistent narrative thread from element-level data to the final criterion interpretation. The qualitative interpretation I_d for dimension d of a given criterion is defined as:

$$I_d = g(L_d) \quad (3)$$

Where:

- I_d : The qualitative descriptor (narrative) assigned to criterion for dimension d.
- g: The mapping function that associates the discrete level L_d (from Equation 2) with its corresponding predefined description in Table 3.

Table 3. Qualitative Interpretation Mapping for Criterion-Level Performance

Dim	L_d	Specific Criterion Interpretation I_d
E1	4	*All of the elements of the criterion are available*
	2	*Most of the elements of the criterion are available*
	1	*There are no available elements or few available elements*
E2	4	*The elements are applied at perfect level*
	3	*The elements are applied at good level*
	2	*The elements are applied at low level*
	1	*The elements are not applied or applied at very low level*
E3	4	*Applied regularly*
	2	*Applied irregularly*
	1	*Rarely applied*
E4	4	*There is a regular and effective assessment*
	3	*There is a regular assessment*
	2	*There is assessment but it is irregular*
	1	*There is no assessment*
E5	4	*Sufficient and varied Evidence is Available*
	3	*Sufficient Evidence is Available*
	2	*Insufficient Evidence is Available*
	1	*No Evidence is Available*
E7	4	*There are regular procedures for improvement and higher results*
	3	*There are regular improvement procedures and good results*
	2	*There may be some limited improvement procedures*
	1	*No improvement procedures are in Place*

5.4.4. Matrix 1: E6 Composite Assessment of Consistency and Evidence

The composite dimension E6 (Combined Consistency and Evidence) provides a higher-order assessment of how regularly a criterion is applied and assessed. The framework processes the interpretations of dimensions E3, E4, and E5 by constructing a Binary Evaluation Vector (V_{in_E6}),

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

which is validated against the matrix columns using a prioritized logical sequence:

$$\boldsymbol{A_{E6}} = \sum (\boldsymbol{\Phi}(\boldsymbol{V_{in\_E6}}, \boldsymbol{B_j})) \quad \geq \quad \boldsymbol{T_{E6}} \qquad (4)$$

Where:

- A_{E6} : The resulting operational level for the composite dimension of regularity and evidence.
- V_{in_E6} : The Binary Evaluation Vector. It is defined as the concatenation of three binary sub-vectors, each representing the qualitative standing of dimensions E3 (Application Regularity), E4 (Assessment Regularity), and E5 (Evidence Availability).
- B_j: The Case Vector for column j, defining the required binary state for each performance scenario (Cases A–E).
- Φ: The Logical Operator (AND/OR) applied to validate the relationship between V_{in_E6} and B_j.
- T_{E6}: The Derived Operational Level, serving as the summation threshold required to confirm a specific E6 status.

Table 4. Matrix 1: E6 Operational Logic and Benchmarking

Dimension	Performance case	A	B	C	D	E
	E6 Compliance level	1	2	3	4	3
	Logical Operator	**OR**	**OR**	**AND**	**AND**	**OR**
	Derived Operational Level (T_{E6})	**1**	**1**	**3**	**3**	**1**
E3	*Applied regularly*	0	0	1	1	0
	Applied irregularly	0	1	0	0	0
	Rarely applied	1	0	0	0	0
E4	*There is a regular and effective assessment*	0	0	0	1	1
	There is a regular assessment	0	0	1	0	0
	There is assessment but it is irregular	0	1	0	0	0
	There is no assessment	0	1	0	0	0
E5	*Sufficient and varied Evidence is Available*	0	0	0	1	1
	Sufficient Evidence is Available	0	0	1	0	0
	Insufficient Evidence is Available	0	1	0	0	0
	No Evidence is Available	1	0	0	0	0

This resulting composite vector is processed through the decision matrix to determine the E6 level using the following prioritized sequence:

- Case A - Level 1: if ($A_{E6} \geq 1$)
- Case B - Level 2: if ($A_{E6} \geq 1$)
- Case C - Level 3: if ($A_{E6} \geq 3$)
- Case D - Level 4: if ($A_{E6} \geq 3$)
- Case E - Level 3: if ($A_{E6} \geq 1$)

5.4.5. *Formula 4: E6 Qualitative Narrative Synthesis*

After the logical validation in Matrix 1 determines the E6 level, the framework generates a Qualitative Narrative Synthesis (I_6). This formula acts as a concatenation engine that combines the individual interpretations of regularity and evidence into a single, human-readable statement. This synthesis is vital for automated reporting, as it provides a clear, transparent justification for the assigned E6 level to stakeholders and auditors. The synthesis of the composite narrative is defined as:

$$\boldsymbol{I_6} = \boldsymbol{I_3} \oplus \boldsymbol{I_4} \oplus \boldsymbol{I_5} \qquad (5)$$

Where:

- I_6: The final composite qualitative narrative for E6.
- ⊕: The concatenation operator used to link the individual dimension narratives into a logically structured sentence.
- I_3, I_4, I_5: The specific qualitative interpretations for Application, Assessment, and Evidence (from Table 4).

5.4.6. *Matrix 2: Final Criterion Compliance Assessment*

The final evaluative judgment for a specific criterion is determined by Matrix 2, which integrates the qualitative interpretations of all six primary dimensions. The framework processes these interpretations by constructing a Binary Evaluation Vector (V_{in}), which is then validated against the matrix columns using a prioritized logical sequence:

$$\boldsymbol{A_C} = \sum (\boldsymbol{\Phi}(\boldsymbol{V_{in}}, \boldsymbol{B_j})) \;\geq\; \boldsymbol{T_{Final}} \qquad (6)$$

Where:

- A_C: The final compliance achievement level for the assessed criterion.
- V_{in}: The Binary Evaluation Vector. It is defined as the concatenation of six binary sub-vectors, each representing the qualitative standing of a specific dimension (E1, E2, E3, E4, E5, E7). For each sub-vector, a value of 1 is assigned to the position corresponding to the active qualitative interpretation (as defined in Table 5), while all others are assigned 0.
- B_j: The Case Vector for column j, defining the required binary state for each performance scenario (Cases A–E).
- Φ: The Logical Operator (AND/OR) applied to validate the relationship between V_{in} and B_j.
- T_{Final}: The Derived Operational Level, serving as the summation threshold required to confirm a specific compliance level.

This resulting composite vector is processed through the decision matrix to determine the final compliance level using the following prioritized sequence:

- Case A - Level 1: if ($A_C \geq 1$)
- Case B - Level 2: if ($A_C \geq 1$)
- Case C - Level 3: if ($A_C \geq 6$)
- Case D - Level 4: if ($A_C \geq 6$)
- Case E - Level 3: if ($A_C \geq 1$)

Table 5.Matrix 2: Final Criterion Compliance Logic

Dimension	Performance case	A	B	C	D	E
	Final Compliance Level	1	2	3	4	3
	Logical Operator	**OR**	**OR**	**AND**	**AND**	**OR**
	Derived Operational Level (T_{Final})	**1**	**1**	**6**	**6**	**1**
E1	*All of the elements of the criterion are available*	0	0	1	1	0
	Most of the elements of the criterion are available	0	1	0	0	0
	There are no available elements or few available elements	1	0	0	0	0
E2	*The elements are applied at perfect level*	0	0	0	1	1
	The elements are applied at good level	0	0	1	0	0
	The elements are applied at low level	0	1	0	0	0
	The elements are not applied or applied at very low level	1	0	0	0	0
E3	*Applied regularly*	0	0	1	1	0
	Applied irregularly	0	1	0	0	0
	Rarely applied	1	0	0	0	0
E4	*There is a regular and effective assessment*	0	0	0	1	1
	There is a regular assessment	0	0	1	0	0
	There is assessment but it is irregular	0	1	0	0	0
	There is no assessment	1	0	0	0	0
E5	*Sufficient and varied Evidence is Available*	0	0	0	1	1
	Sufficient Evidence is Available	0	0	1	0	0
	Insufficient Evidence is Available	0	1	0	0	0
	No Evidence is Available	1	0	0	0	0
E7	*There are regular procedures for improvement and higher results*	0	0	0	1	1
	There are regular improvement procedures and good results	0	0	1	0	0
	There may be some limited improvement procedures	0	1	0	0	0
	No improvement procedures are in Place	1	0	0	0	0

5.4.7.Formula 5: Final Criterion Interpretation

The final stage of the assessment process is the generation of a Comprehensive Narrative Statement (I_{Final}). While Matrix 2 provides the numerical compliance level, Formula 5 synthesizes the qualitative interpretations from all six primary dimensions into a cohesive, human-readable summary. This automated narrative ensures that stakeholders receive not only a score but also a transparent, descriptive justification for the criterion's final standing. The comprehensive narrative synthesis is defined as:

$$\boldsymbol{I_{Final} = I_1 \oplus I_2 \oplus I_3 \oplus I_4 \oplus I_5 \oplus I_7} \quad (7)$$

Where:

- I_{Final} : The complete human-readable assessment statement for the criterion.
- $\oplus$: The concatenation operator used to integrate the individual dimensional interpretations into a logical sequence.
- I_d : The qualitative interpretation for each dimension d (from Table 5).

5.4.8.Formula 6: Global Standard Aggregation and Compliance Constraints

The final step in the framework is the calculation of the Global Standard Level (***GSL***). This aggregation does not rely on a simple average; instead, it enforces a "minimum quality floor" by applying logic-based constraints. These constraints ensure that a standard cannot be considered "Proficient" or "Exemplary" if fundamental practices are missing. The baseline aggregation is defined by the average of all criteria scores, rounded to the nearest integer:

$$\boldsymbol{GSL_{base} = round(\frac{1}{N}\sum_{i=1}^{N} A_{C,i})} \quad (8)$$

To maintain the integrity of the assessment, the final level is subject to a conditional capping function. If critical thresholds are not met, the ***GSL*** is automatically downgraded to Level 2 (Developing), regardless of the mathematical average:

$$\boldsymbol{GSL_{Final}} = \begin{cases} \boldsymbol{\min(2, GSL_{base})} & \boldsymbol{if \exists A_{C,i} < 3 \ for\ i \in Essential\ Crit} \\ \boldsymbol{\min(2, GSL_{base})} & \boldsymbol{if Count(A_{C,i} < 3) > 2\ for\ i \in NonEssent\ Crit} \\ \boldsymbol{GSL_{base}} & \boldsymbol{otherwise} \end{cases} \quad (9)$$

Where:

- N: The total number of criteria within the standard.
- $A_{C,i}$: The final achievement level of criterion i (calculated in Formula 6).
- Essential Criteria Violation: An automatic cap is triggered if any criterion marked as "Essential" falls below Level 3.
- Non-Essential Criteria Tolerance: A cap is triggered if more than two "Non-Essential" criteria fall below Level 3.

Formula 6 operationalizes the Essential Criteria Constraint within the standard aggregation logic. By applying this constraint, the ASEP system identifies critical deficiencies immediately; if the binary vector V_{in} fails the Essentiality check, the criterion level is automatically capped at Level 1 or 2, regardless of performance in non-essential dimensions. This prevents the 'inflation of quality' where peripheral successes mask core compliance gaps, ensuring that the final level (L_f) reflects absolute adherence to mandatory standards rather than a numerical average. This mathematical safeguard ensures that institutional excellence is grounded in foundational integrity, preventing a program from achieving a high accreditation status while missing fundamental quality pillars.

5.5. WORKED EXAMPLE: COMPREHENSIVE EVALUATION OF CRITERION 1.1.1

To demonstrate the framework's operational utility, Criterion 1.1.1—"The program's mission and goals are consistent with the mission of the institution and guide all its operations"—is evaluated. This criterion is broken down into five constituent elements as shown in Table 6.

5.5.1. ELEMENT-LEVEL ASSESSMENT

Initial scoring is conducted at the element level across the six dimensions. These raw scores represent the degree of compliance for each specific component of the mission-alignment process.

Table 6.Element-Level Scores for Criterion 1.1.1

Element	E1	E2	E3	E4	E5	E7
Alignment with university mission	4	4	4	4	4	4
Alignment with college mission	0	3	2	3	3	3
Alignment of goals with university	4	2	1	2	2	2
Alignment with college goals	4	1	1	1	1	1
Mission guides operations	4	4	4	4	4	4

5.5.2. DIMENSIONAL ANALYSIS AND QUALITATIVE MAPPING

By applying Formulas 1 and 2, the raw element scores are aggregated into dimensional percentages and mapped to their corresponding qualitative levels (Table 7). Dimensional percentages are computed using Formula 1, where $n = 5$ elements. For example, E1 scores are [4, 0, 4, 4, 4], yielding $(4 + 0 + 4 + 4 + 4)/(4 \times 5) = 16/20 = 80\%$. Similarly, E7 scores [4, 3, 2, 1, 4] sum to 14, giving $14/20 = 70\%$. All percentages are then mapped to qualitative levels using the thresholds in Table 2.

Table 7. Dimensional Analysis for Criterion 1.1.1

Dimension	%	Level	Qualitative Description
E1	80%	2	*Most of the elements of the criterion are available*
E2	70%	3	*The elements are applied at good level*
E3	60%	2	*Applied irregularly*
E4	70%	3	*There is a regular assessment*
E5	70%	3	*Sufficient Evidence is Available*
E7	70%	3	*There are regular improvement procedures and good results*

This analysis reveals that while most elements are present (E1 = Level 2) and implementation quality is adequate (E2 = Level 3), the program suffers from inconsistent application (E3 = Level 2) and lacks systematic improvement mechanisms (E7 = Level 3). These findings directly inform targeted interventions, such as establishing standardized implementation protocols and enhancing faculty training on continuous improvement cycles.

5.5.3. RULE-BASED MATRIX APPLICATION

The dimensional profile is then processed through the two-stage matrix logic to determine the final standing.

- **Matrix 1 (E6 Composite):**
  - Input: E3=" Applied irregularly", E4=" There is a regular assessment", E5=" Sufficient Evidence is Available"
  - V_{in_E6}: [E3=(0,1,0), E4=(0,1,0,0), E5=(0,1,0,0)]
  - A_{E6} Calculation for all Bj: [0, 1, 2, 0, 0] to be compared with [1, 1, 3, 3, 1]
  - Result: A_{E6} [2] ≥ 1 → E6 Compliance level = 2
- **Matrix 2 (Final Criterion Level):**
  - Input: All six-dimensional descriptions from Table 7
  - Vin: [E1=(0,1,0), E2=(0,1,0,0), E3=(0,1,0), E4=(0,1,0,0), E5=(0,1,0,0), E7=(0,1,0)]
  - A_C Calculation for all Bj: [0, 3, 2, 0, 0] to be compared with [1, 1, 6, 6, 1]
  - Result: A_C [2] ≥ 1 → Final Compliance Level of the Criterion 1.1.1= 2

5.5.4. GLOBAL STANDARD AGGREGATION AND STRATEGIC DIAGNOSTIC

Finally, the system aggregates multiple criteria to determine the Global Standard Level (*GSL*). Example Scenario: Consider a standard composed of 8 criteria with the following Compliance Level: [2, 3, 3, 4, 2, 3, 3, 4].

- Mathematical Average: 3.0
- Constraint Check: Criterion 1.1.1 is designated as Essential. Because its score is L < 3, equation 9 triggers an automatic cap.
- Final Decision: The Standard is awarded Level 2 (Developing), despite the higher average, enforcing compliance with essential quality pillars.

5.5.5. STRATEGIC VALUE AND DIAGNOSTIC PRECISION

This framework transforms a simple score into a powerful diagnostic tool. By maintaining dimensional separation, the system identifies that for Criterion 1.1.1:

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

- Strengths: Quality of implementation (E2), Assessment (E4), and Improvement (E7) are all performing at Level 3
- Critical Gaps: The primary failures reside in Element Availability (E1) and Regularity (E3)
- Actionable Intelligence: Strategic resource allocation should prioritize foundational implementation over redundant assessment refinement; instead, they must be focused on ensuring all mission-alignment elements are formally established and applied consistently across all college operations.

The mathematical rigor of the "Sum-and-Compare" logic ensures 100% reproducibility, providing a transparent roadmap for institutional improvement that aligns perfectly with international accreditation standards. Thus, the framework functions not only as an evaluator but as a prescriptive diagnostic tool, program leaders know exactly whether to focus on policy creation (E1), staff training (E2/E3), documentation systems (E5), or feedback loops (E7).

6. IMPLEMENTATION AND CASE STUDY: STRATEGIC IMPACT

6.1. System Architecture: The ASEP Platform

The proposed quantitative framework was operationalized through the Accreditation Standards Evaluation Platform (ASEP)—not merely a digital tool but a methodological enforcer that embeds the rule-based logic directly into the user interface. ASEP requires structured inputs and blocks submission unless evidence is linked and dimensional scores comply with logical constraints, thereby transforming self-evaluation from a narrative exercise into a verifiable, auditable process. Developed using Microsoft Excel 365 and hosted on Northern Border University's OneDrive infrastructure, the platform provides a centralized environment for the complete NCAAA accreditation lifecycle. By integrating real-time data synchronization with cloud-based collaboration, ASEP serves as a unified "single source of truth" for all academic programs during the self-evaluation process.

6.2. Integrated Platform Interfaces

The ASEP architecture is comprised of five specialized interfaces designed to bridge the gap between raw data and institutional decision-making.

6.2.1. Institutional and Compliance Dashboards

The Institutional Dashboard (Figure 1) provides program leadership with a high-level monitoring hub, offering real-time visibility into overall completion percentages and evidence availability rates. This macro-level view is complemented by the Standard Compliance Dashboard (Figure 4), which integrates the quantitative results into the institutional governance workflow. This interface bridges the gap between assessment and action by hosting the review workflow, where the Permanent Assessment Committee and the Deanship of Quality provide formal recommendations and track institutional response.

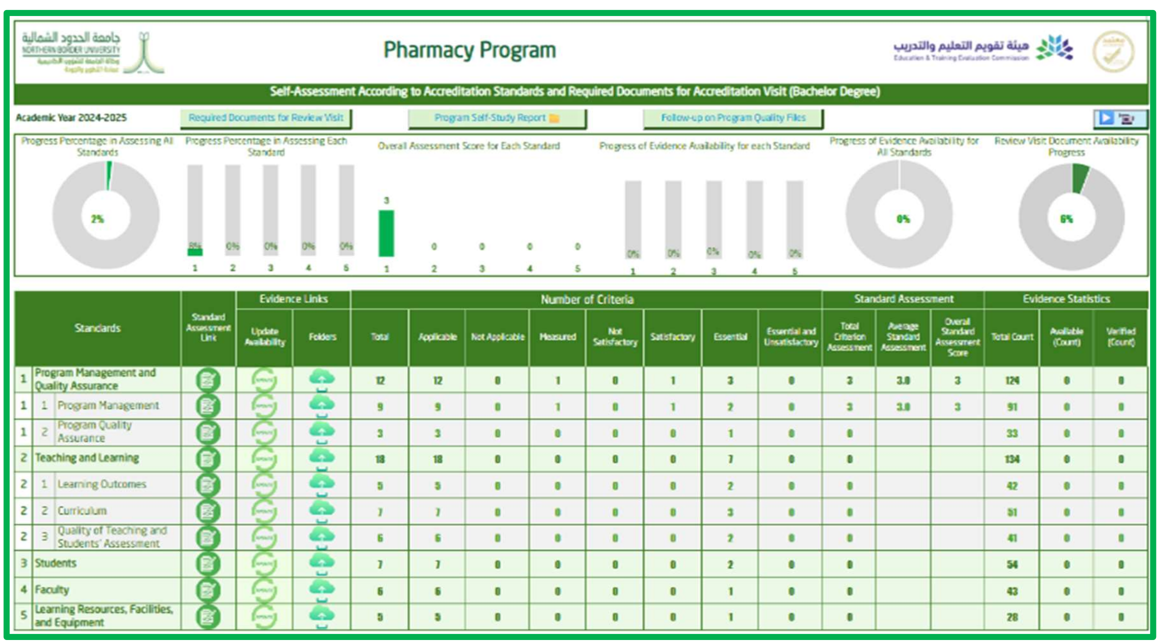

Figure 1. Dashboard Showing Program Self-Assessment Progress

6.2.2. EVIDENCE AND DOCUMENT MANAGEMENT

The Evidence Management System (Figure 2) transforms evidence collection from an ad-hoc activity into a systematic, auditable procedure. It utilizes a hierarchical structure that links every criterion to predefined identification codes and recommended documentation types. This rigorous tracking extends to the Review Visit Interface (Figure 5), which manages the final documentation lifecycle—including audit, arbitration, and report approval—to ensure total readiness for external review panels.

6.2.3. CRITERION EVALUATION ENGINE

At the heart of the platform is the Criterion Evaluation Interface (Figure 3). This interface operationalizes the mathematical formulas discussed in Section 4 by requiring assessors to provide element-level scores across the seven performance dimensions. By forcing the separation of "Application Quality" from "Assessment Regularity," the system eliminates the ambiguity of holistic subjective judgments and generates the granular data necessary for diagnostic analysis.

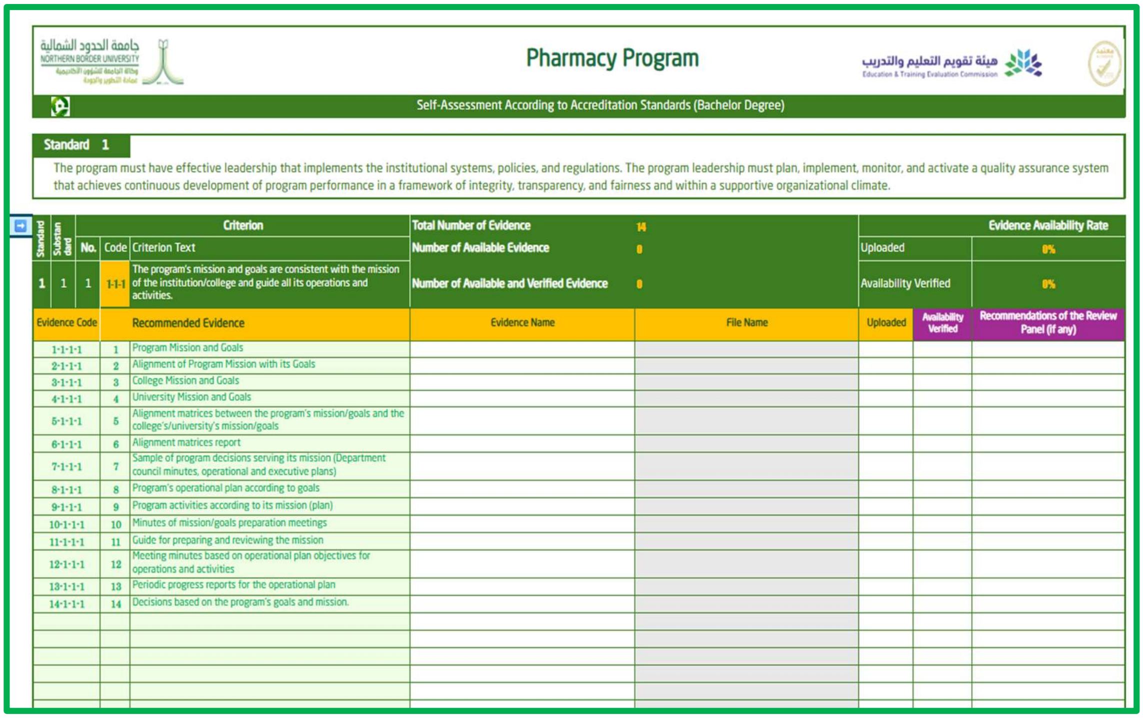

Figure 2. Evidence Management Interface with OneDrive Integration

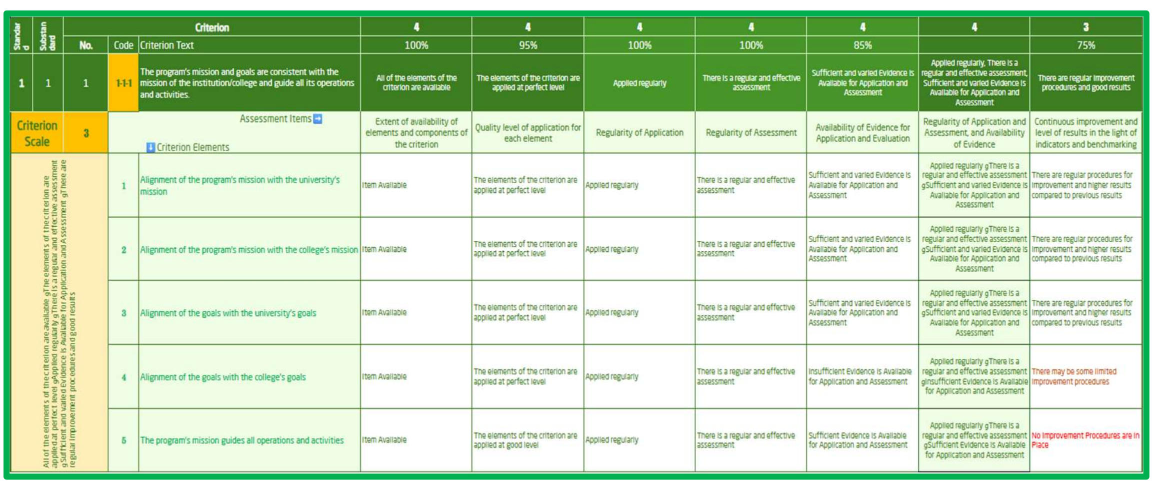

Figure 3. Granular Criterion Evaluation Interface

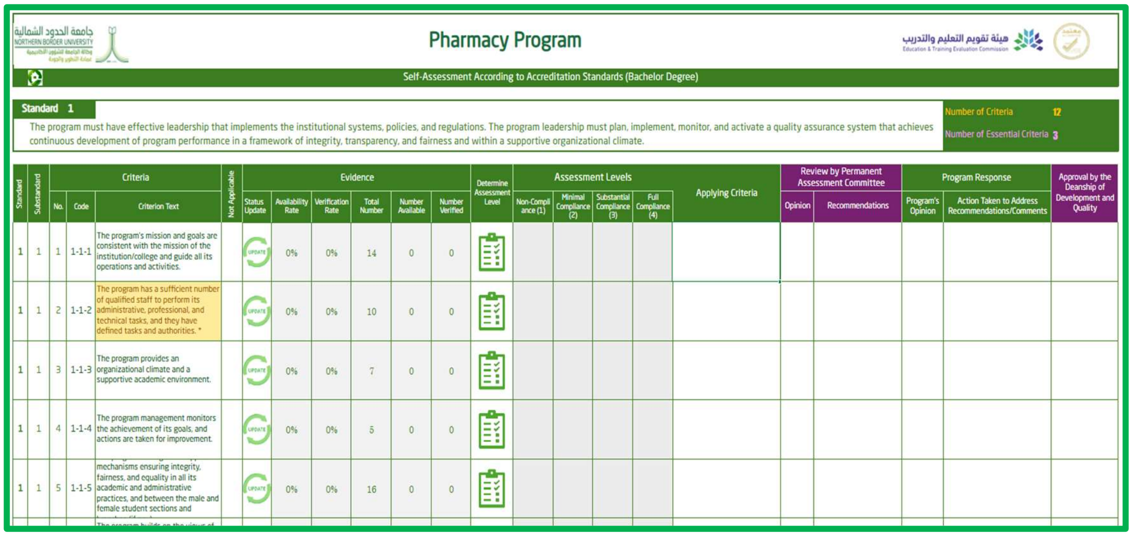

Figure 4. Compliance Monitoring Matrix with Review Workflow

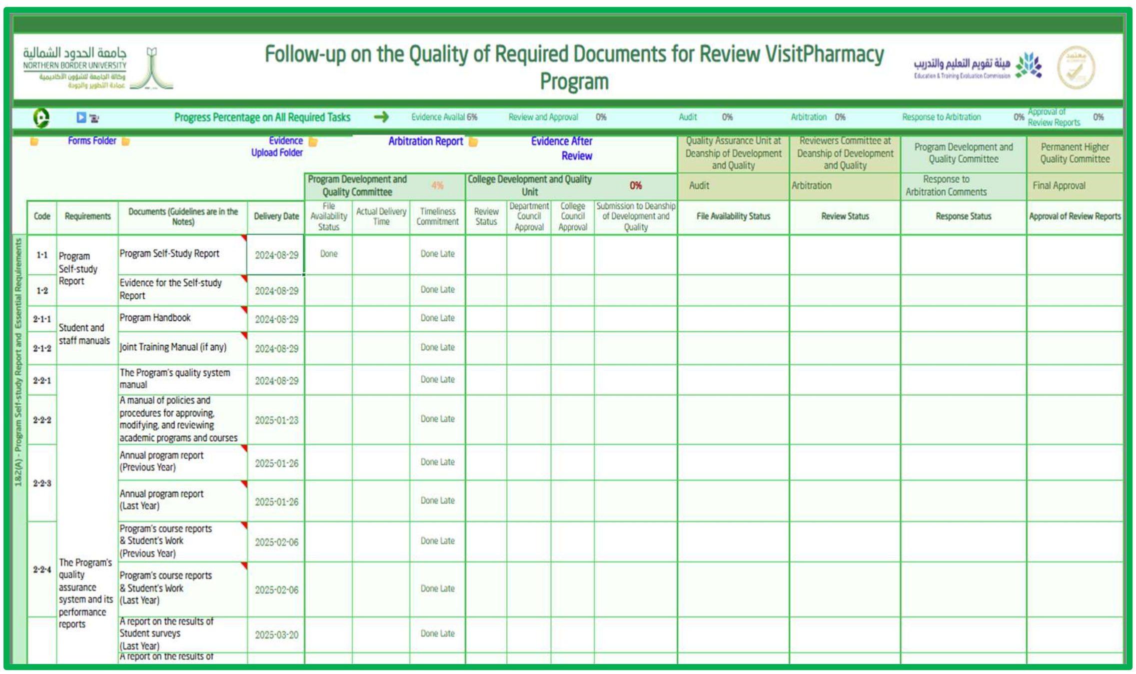

Figure 5.Document Follow-up Interface for Review Visit Requirements

6.3. *VALIDATION METRICS AND COMPARATIVE PERFORMANCE*

To empirically evaluate the effectiveness of the ASEP platform, a longitudinal comparative analysis was conducted between the traditional 2022/2023 evaluation cycle and the last 2024/2025 ASEP-driven cycle. The results, summarized in Table 8, demonstrate significant advancements in both operational efficiency and evaluative realism. Empirical results indicate that the framework significantly streamlined the accreditation workflow, reducing assessment processing time by 60% and decreasing missing evidence rates by 66%. These gains were achieved through mandatory evidence-linking and automated aggregation, which eliminated the administrative bottlenecks of traditional manual reporting. Simultaneously, the 62% average reduction in absolute score discrepancy validates the 'Sum-and-Compare' logic's ability to align decentralized self-assessments with independent audit findings. Furthermore, the 30% reduction in inflated compliance ratings confirms that the framework effectively mitigates 'leniency bias,' replacing subjective optimism with a realistic diagnostic of institutional performance. The observed shift in compliance distribution—specifically the migration of scores from Level 3 down to Level 2—should not be interpreted as a decline in institutional quality. Rather, it represents the 'Correction of Leniency.' The framework's dimensional rigor exposed implementation gaps that were previously hidden behind well-written narratives. This shift transforms the self-evaluation from a promotional document into an honest roadmap for genuine quality improvement.

Table 8. Impact of ASEP on Evaluation Accuracy and Efficiency

Metric	22/23	24/25	Improv.	Rationale
Missing Evidence Rate	35%	12%	⬇ 66%	Mandatory evidence linking with real-time tracking
Score Discrepancy*	2.1 points	0.8 points	⬇ 62%	Objective rule-based scoring eliminates subjective aggregation
Criteria Ass. Level 3	65%	45%	⬇ 30%	More realistic distribution through dimensional assessment
Criteria Ass. Level 2	15%	35%	⬆ 57%	Reduced leniency bias exposes actual weaknesses
Assessment Time per Criterion	45 minutes	18 minutes	⬇ 60%	Automated calculations and structured inputs

* Average absolute difference between program self-assessment scores and internal review committee scores

6.4. *Diagnostic Case Study: Standard 2 (Teaching and Learning)*

The framework's most significant contribution is the transformation of self-evaluation into Actionable Intelligence. A diagnostic analysis of three criteria under Standard 2 reveals how dimensional scores dictate specific development paths.

For Criterion 2-1-4, the profile (4, 3, 2, 2) indicates that while assessment policies are excellent, the program lacks consistency in application. Development efforts should therefore prioritize standardized assessment schedules over further policy writing. The profile for Criterion 2-2-1 (4, 2, 1, 1) serves as a primary example of the Essential Criteria Constraint in action. Despite having a perfect score in policy availability (E1=4), the failure in implementation regularity triggered constraint, capping the overall level at 'Adequate' (Level 2). In a traditional holistic system, the strong policy documentation might have pulled the average score higher, potentially leading to an inaccurate 'Good' (Level 3) rating. However, the ASEP logic ensures that missing foundational implementation results in a realistic diagnostic rating. This logic contributed to the 60% increase in the accurate identification of Level 2 implementation gaps observed across the university, proving that the system effectively forces a more critical and honest appraisal of program weaknesses.

Table 9. Multi-Dimensional Diagnostic of Standard 2 Criteria

Criterion	E1	E2	E3	E4	Level	Program Development Insight
2-1-4: Learning Outcomes Assessment	4	3	2	2	**3**	Prioritize longitudinal assessment consistency over policy redrafting rather than policy development
2-2-1: Curriculum Design	4	2	1	1	**2**	Critical need for implementation support, not additional planning
2-3-5: Assessment Quality	3	4	4	3	**4**	Maintain excellence while documenting improvement processes

6.5. *Efficiency and Reproducibility*

The ASEP framework delivered a 60% gain in evaluation efficiency while ensuring 100% reproducibility; two different evaluators given the same evidence will now arrive at an identical score. While initial challenges included faculty resistance and varying digital literacy, a phased rollout and stakeholder training successfully transformed the system from a compliance burden into a trusted diagnostic tool. This data-driven approach ensures that resource allocation is strategically aligned with actual performance gaps, rather than perceived needs.

7. DISCUSSION AND COMPARATIVE CRITIQUE

7.1 Comparative Analysis with State-of-the-Art

The primary IT contribution of the ASEP framework lies in its transition from 'subjective weighting' to 'deterministic logic.' Unlike the widely used Analytic Hierarchy Process (AHP) described by Saaty [13], which relies on expert-defined pairwise comparisons that can lead to inconsistent results, ASEP utilizes a Binary Evaluation Vector (V_{in}) to ensure objectivity. While traditional AHP models often struggle with an 'Inconsistency Ratio,' the ASEP framework effectively bridged a 62% average absolute discrepancy between self-assessment and audit scores, proving that deterministic rules provide higher reliability than expert-weighting alone.

While Mondal [9] and Naim et al. [10] focus on the automation of KPI tracking, their models often treat all evidence with equal structural weight. In contrast, ASEP introduces the Consistency Layer (E6), which programmatically prevents a standard from being marked as 'Met' if the evidence does not mathematically align with the regularity of application. The impact of this logic is evidenced by the 66% reduction in missing or undocumented evidence observed in our pilot study. This 'sum-and-compare' logic ensures a 100% reproducible outcome, effectively solving the 'black-box' assessment problem found in previous literature.

7.2 Limitations and Shortcomings

Despite its strengths, the current iteration of the ASEP framework has specific shortcomings:

- Manual Input Dependency: Unlike the NLP-based approach suggested by Nguyen and Tran [11] for automated evidence detection, ASEP currently requires human evaluators to manually input the binary status for dimensions E1-E7. However, even with manual entry, the system achieved a 60% reduction in processing time by automating the aggregation and reporting phases.
- Framework Specificity: The logic is currently mapped to NCAAA standards. Extending this to other international frameworks (like ABET or QAA) requires a manual reconfiguration of the underlying rule-set.
- Scalability of Documentation: As the volume of digital evidence grows, the system requires higher cloud-compute resources to maintain the real-time responsiveness of the decision-support dashboards.

8. CONCLUSION AND FUTURE WORK

This study successfully addresses the primary research problem posed in the introduction: the inherent subjective bias and lack of reproducibility in traditional, holistic accreditation models. The results confirm that by formalizing accreditation standards into a structured algorithmic logic, institutional assessment can transition from a heuristic, narrative-driven process to a reproducible, evidence-based computational workflow. This work provides a definitive answer to the research objective by proving that the mathematical coupling of qualitative indicators with quantitative compliance levels—via Binary Evaluation Vectors (V_{in})—effectively eliminates the "black-box" variability of human judgment.

The central argument of this research is that institutional quality is a data-integrity challenge rather than a purely interpretive exercise. The empirical validation of the ASEP framework (as detailed in Table 8) supports this insight, yielding a 60% reduction in assessment processing time and a 66% reduction in missing or undocumented evidence. Crucially, the system effectively bridged the 62% average absolute discrepancy previously observed between self-assessment and audit scores, while correcting inflated compliance ratings by 30%. These findings validate the hypothesis that deterministic logic provides a higher-fidelity diagnostic of institutional excellence than subjective human aggregation.

While the current implementation focuses on NCAAA standards, the underlying system architecture is domain-agnostic. The core intellectual contribution of this work is the "Logic Layer"—the modular design of decision matrices that ensures every compliance decision is traceable, auditable, and mathematically rigorous. This establishes the framework not merely as an accreditation tool, but as a scalable model for automated compliance monitoring in any sector requiring rigorous auditing.

While this study establishes a foundation for deterministic quality assessment, several open research issues remain for the academic community to explore:

1. AI-Driven Evidence Verification: A primary pending issue is the automation of the evidence-mapping process. Future research should explore integrating Natural Language Processing (NLP) to automate the validation of uploaded documents, addressing the current challenge of manual verification overhead.
2. Predictive Analytics: An open question remains regarding how longitudinal data within ASEP can be leveraged. Developing Machine Learning models capable of identifying performance trends and predicting accreditation risks is a critical next step for data-driven quality management.
3. Cloud-Native SaaS Migration: Transitioning to a fully web-based SaaS model with API interoperability remains a technical frontier required for seamless integration with existing Institutional Research (IR) and Learning Management Systems (LMS).
4. Cross-Framework Adaptability: A significant opportunity for future research lies in developing a dynamic "Translation Layer." This would allow the ASEP deterministic rules to be automatically remapped to various international standards (e.g., ABET, QAA), resolving the current constraint of framework-specific configuration.

ACKNOWLEDGMENT

The authors extend their appreciation to the Deanship of Scientific Research at Northern Border University, Arar, KSA, for funding this research work through project number NBU-FFR-2026-417-

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

02. The authors would also like to thank the Deanship of Development and Quality at Northern Border University for providing the institutional support and data access necessary to implement and validate the NBU Quality Framework.

REFERENCES

[1] Alqahtani, A.Y., Makki, A.A. and Abdulaal, R.M.S. (2023), ‘A proposed NCAAA-based approach to the self-evaluation of higher education programs for academic accreditation: a comparative study using TOPSIS’, Decision Science Letters, 12(3), pp. 333–346.

[2] Chang, D.Y. (2010) ‘Fuzzy decision making in education quality evaluation’, Applied Soft Computing, 10(4), pp. 1037–1048.

[3] Dawson, P., Bennett, S. and Bearman, M. (2022) ‘Can assessment and accreditation systems support authentic quality improvement in higher education?’, Assessment & Evaluation in Higher Education, 47(4), pp. 525–537.

[4] Duarte, N. and Vardasca, R. (2023) ‘Literature review of accreditation systems in higher education’, Education Sciences, 13(6), pp. 582.

[5] Education and Training Evaluation Commission (ETEC) (2021) Standards for quality assurance and accreditation of higher education institutions. Available at: https://www.etec.gov.sa (accessed 10 June 2025).

[6] European Association for Quality Assurance in Higher Education (2015) Standards and Guidelines for Quality Assurance in the European Higher Education Area (ESG). 3rd ed. Brussels: ENQA.

[7] Garcia, P. and Ramirez, L. (2021) ‘Designing KPI-based dashboards for academic accreditation management’, Computers and Education: Artificial Intelligence, 2, pp. 100023.

[8] Harvey, L. and Williams, J. (2010) ‘Fifteen years of quality in higher education: much ado about nothing?’, Quality in Higher Education, 16(1), pp. 5–21.

[9] Mondal, S.R. (2025) ‘Automating KPI measurement: a sustainable solution for educational accreditation’, Sustainability, 17(5), pp. 1968.

[10] Naim, A., Alnfiai, M.M. and Almalki, N.S. (2024) ‘Information systems based working model for the assessment of program learning outcomes in measuring the quality in higher education’, Research Square. Available at: https://doi.org/10.21203/rs.3.rs-4123456/v1 (accessed 10 June 2025).

[11] Nguyen, L. and Tran, M. (2020) ‘Automating evidence detection in accreditation reports using natural language processing’, IEEE Access, 8, pp. 98763–98775.

[12] Quality Assurance Agency for Higher Education (2018) The UK quality code for higher education. Available at: https://www.qaa.ac.uk (accessed 10 June 2025).

[13] Saaty, T.L. (2008) ‘Decision making with the analytic hierarchy process’, International Journal of Services Sciences, 1(1), pp. 83–98.

[14] Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Cambridge, MA: Harvard University Press.

[15] Wen, C., Wang, S. and Lu, Y. (2020) ‘Multi-criteria evaluation of higher education quality using analytic hierarchy process (AHP)’, Studies in Educational Evaluation, 67, pp. 100–115.

[16] ABET (2023) Criteria for accrediting engineering programs, 2023–2024. Available at: https://www.abet.org (accessed 10 June 2025).