National Commission for Certifying Agencies Guidance Document: Use of Artificial Intelligence in Certification Programs

Version 1.1

Released April 2, 2025

Table of Contents

NCCA Guidance: Use of Artificial Intelligence in Certification Programs

The National Commission for Certifying Agencies (NCCA) developed this guidance to provide certification programs with a framework for maintaining compliance with NCCA Standards when integrating Artificial Intelligence (AI) into certification program activities.

For the purposes of this guidance, references to AI include both AI systems and Generative AI[1]. The following definitions apply:

> **AI system:** A machine-based system that processes input data, recognizes patterns, generates outputs, and adapts its responses based on encoded knowledge or learning. AI systems range from rules-based expert systems and machine-learning models to more advanced applications such as large language models and intelligent agents.
>
> **Generative AI:** A subset of AI systems that produces new content, predictions, or outputs, rather than simply analyzing or classifying existing data. Unlike traditional AI models that rely on rule-based decision-making or predefined outputs, Generative AI creates text, images, code, or other data representations based on training. This definition includes expert systems, machine-learning applications, large language models, intelligent agents, and similar systems.

AI has the potential to inform core certification or test-development decision-making.[2] The intent of this guidance is not to restrict innovation but rather to ensure that AI-driven solutions align with the NCCA Standards and uphold the validity and reliability of certification decisions.

This guidance document is considered an update to the Commentary of the 2021 NCCA Standards, and the NCCA may base accreditation decisions on findings of noncompliance with the Standards and Essential Elements, as further elaborated in the Commentary and this guidance document.

---

[1] References to AI exclude traditional computer adaptive testing (CAT), multi-stage testing (MST), linear on-the-fly testing (LOFT), and automated test assembly (ATA) algorithms, as these systems rely on predefined rules even if they exhibit adaptiveness and randomness.

[2] As per 2021 NCCA Standards, Standard 2, Commentaries 2-3:

"2.) Essential certification decisions refer to the core aspects of a certification program, such as eligibility standards; standards for initial certification and maintaining certification; disciplinary determinations; the development, administration, and scoring of examinations; and the selection of subject-matter experts (SMEs).

3.) Decisions that are NOT considered essential include those decisions related to employee selection, office location, marketing and communications efforts, and final budget or contract approval as long as sufficient financial resources are provided for the certification program and policies and procedures are in place to provide for autonomy in essential certification decisions."

Maintaining Accreditation and Adapting to Evolving AI Standards

This guidance provides an initial framework for AI use in certification programs. However, these considerations are not exhaustive, and programs remain responsible for ensuring that AI use aligns with all relevant Essential Elements of the NCCA Standards.

As AI technologies and regulatory landscapes evolve, the NCCA may amend this guidance as needed to reflect emerging developments and ensure continued alignment with accreditation requirements.

Human Oversight of AI in Certification Programs

While AI may assist in certification activities, AI systems are not a substitute for qualified human personnel in core certification or test-development decisions. AI must not operate independently in these functions without appropriate human oversight. Certification programs using AI must ensure that all AI-assisted decisions are validated by qualified human personnel and comply with all applicable legal and accreditation requirements. AI systems may perform useful work, such as identifying trends, streamlining processes, or adding a level of quality assurance screening of output that is originally generated by human personnel, but they cannot serve as the final authority in any function that requires professional judgment, ethical considerations, or regulatory compliance.

To maintain integrity in AI-assisted processes, certification programs must ensure that AI systems are subject to qualified human personnel oversight, similar to the requirements for staff and vendors under NCCA standards.

Standard 2: Governance and Autonomy

Certification programs must be overseen by an autonomous certification board or commission of qualified human personnel. AI systems may be used to support governance; however, AI systems alone cannot hold final decision-making authority related to policies, governance, or certification approvals.

Standard 3: Education, Training, and Certification

Certification programs must maintain a clear separation between AI used for creating, endorsing, or delivering preparatory materials and AI involved in essential certification decisions (See Standard 3, Commentary 2). With or without AI, a firewall must be maintained between education/training and certification.

AI may assist in the development of education and training materials, but its use of accredited program materials must be restricted to:

- Publicly available resources, including exam blueprints, competency frameworks, and published guidelines.
- Authorized resources and non-secure reference materials explicitly designated for learning purposes, including proprietary textbooks, or training materials, to the extent consistent with copyright law protections.

AI used in training and preparatory content must **not** be trained on or have access to:

- Confidential examination content, including non-public test items, scoring rubrics, proprietary examination specifications.
- Secure testing materials or any information that could compromise exam validity.

Programs using AI in education or training must:

- Document AI use.
- Maintain policies that prohibit AI from accessing or learning from secure test materials.
- Establish procedures for human oversight of AI-generated educational and exam-related training content.

Certification programs that integrate AI into education or training functions must disclose their AI use to the NCCA and maintain clear documentation demonstrating compliance with NCCA Standard 3. While public disclosure is not required, programs must ensure transparency in their processes and be prepared to provide evidence of AI integration, oversight, and adherence to accreditation requirements upon request.

Standard 5: Human Resources

Certification programs must ensure qualified human personnel oversee the AI systems to maintain integrity in AI-assisted processes. Human personnel are responsible for:

- Understanding the AI system's capabilities and limitations.
- Interpreting and appropriately applying AI-generated outputs.
- Identifying and mitigating potential biases or anomalies.
- Overriding AI-generated recommendations when necessary.

All types of AI-assisted decisions—whether favorable or adverse—must be audited to ensure that the system functions as intended. However, this audit process does not require 100% real-time monitoring of every decision. Instead, programs must implement a structured and periodic process that evaluates:

- AI-generated decisions for accuracy, fairness, and consistency at appropriate intervals.
- Identifies potential biases, errors, or unintended patterns that could impact decision outcomes.

- Incorporates a random or targeted sampling of AI-assisted decisions to verify compliance with policies and accreditation standards.
- Ensures that adverse decisions (See Standard 6, Commentary 5) receive a human-led review before finalization, allowing for intervention when necessary.

Standard 6: Information for Candidates

Programs that use AI-driven chatbots[3] or automated systems for communication with candidates, certificants, and interested parties must ensure transparency, accuracy, and accessibility of information. Interested parties should be informed when AI-generated responses are used and provided with a clear disclaimer stating that AI responses are for general informational purposes only and do not constitute official certification decisions.

To mitigate risks associated with misinformation, programs should establish a process for intervening in service disputes regarding AI inaccuracies. If an interested party disputes an AI-provided response, qualified human personnel must be available to provide clarification or make corrections as necessary. Additionally, programs should implement a structured feedback mechanism where users can document their complaint, provide rationale, and request human review. This feedback should be systematically reviewed by human personnel, analyzed, and used to refine AI outputs over time.

If functioning as intended, programs do not need to audit every AI-generated response but should implement sample-based reviews and error-tracking mechanisms to identify trends and ensure quality control.

Standard 7: Program Policies

Certification programs utilizing AI systems must "comply with applicable laws and regulations" (See Standard 7, Commentary 8) The NCCA Standards do not advise programs as to specific regulatory requirements that must be followed. It is the responsibility of accredited programs to remain informed about laws and case decisions governing AI use in their operational jurisdictions and to document their compliance with policies and procedures accordingly.

Additionally, programs must ensure that AI does not make final determinations in adverse certification decisions.

[3] AI-driven chatbots are machine-based conversational systems that use AI and Natural Language Processing (NLP) to engage with users, answer questions, and provide information in real time. These chatbots can be rule-based, following predefined scripts, or adaptive, leveraging machine learning to generate dynamic responses based on user input.

Standard 8: Awarding of Certification

AI systems cannot independently determine a candidate's certification status. Certification decisions must be based on an examination that meets applicable NCCA Standards and must remain under qualified human personnel oversight. While AI may assist in evaluating candidate performance, programs must ensure that all AI-assisted assessments comply with accreditation requirements and maintain transparency, validity, and fairness.

Any AI-generated outputs contributing to certification decisions must be reviewed, audited, and approved by qualified human personnel before final determinations are made. Programs must implement safeguards to prevent AI from making final, independent certification decisions, including processes for human review and intervention. Additionally, AI-generated decisions must be documented and auditable.

Programs utilizing AI in candidate evaluations must establish mechanisms for bias mitigation, transparency, and ongoing validation to uphold fairness and reliability.

Standard 9: Records Retention and Management Policies

Certification programs must ensure that all records related to candidates, examinations, and certification activities are securely stored and retained in compliance with NCCA Standards and applicable legal requirements.

If AI is used to assist in records retention or document management, programs must ensure that AI-driven systems maintain data integrity, auditability, and compliance with retention policies. AI must not be trained on or modify confidential candidate or examination records unless explicitly authorized by program policies and unless safeguards to preserve confidentiality are in place. Programs must document how AI interacts with records and ensure that qualified human personnel oversee AI-assisted record-keeping processes to prevent inaccuracies, unauthorized deletions, or breaches of confidentiality.

Certification programs should also have clear policies for responding to potential data breaches or misuse of AI-stored records. If AI-generated data summaries or analyses are used for reporting, programs must ensure that AI-derived outputs do not compromise the security or confidentiality of retained records.

Standard 10: Confidentiality

Certification programs must ensure that AI systems handling confidential candidate, exam content and data, and certification information comply with all applicable confidentiality policies, NCCA Standards, and legal requirements. AI systems must not be trained on, store, or process confidential data in ways that could compromise exam security, candidate privacy, or certification

integrity. This includes implementing safeguards to prevent confidential information from being incorporated into AI training data, prompts, or external AI models that extend beyond authorized personnel.

Any AI system used to process or analyze confidential certification data must be restricted to personnel with designated access rights, including staff, Subject Matter Experts (SMEs), and approved vendors. Programs must audit AI interactions with confidential data to ensure compliance with security and accreditation requirements.

If AI-generated insights or summaries are used in certification activities, programs must establish clear verification processes to ensure that AI does not expose, or misuse protected information. Additionally, certification bodies should have policies for remediating unintended AI breaches, including corrective actions, reporting obligations, and human intervention mechanisms to prevent further exposure.

To maintain compliance, programs must periodically review and update confidentiality policies as AI capabilities evolve. Documentation must be maintained to demonstrate how AI aligns with confidentiality requirements.

Standard 11: Conflict of Interest

AI systems do not inherently possess personal, financial, or organizational conflicts of interest, as they lack independent decision-making motives. However, implementing and managing AI within a certification program can introduce risks that must be mitigated to maintain compliance with NCCA Standards. Programs must ensure that AI-driven processes do not unintentionally create conflicts of interest by influencing certification decisions in ways that compromise fairness, objectivity, or transparency.

To manage potential conflicts, AI systems must not be used in ways that prioritize specific candidate groups, vendors, or internal stakeholders without justification based on objective, documented criteria. Additionally, AI models must not be trained on proprietary or confidential data that could create undue advantages for any group or individual involved in the certification process.

The program must conduct thorough due diligence in selecting an AI system for certification activities and certification bodies remain ultimately accountable for any vendor-mediated AI usage. Regular audits should be conducted to assess AI's role in certification governance and decision-making. If an AI system is found to introduce biases or unfair advantages, corrective measures must be implemented.

Standard 12: Security

Certification programs must establish and maintain comprehensive security policies and procedures to protect the integrity, confidentiality, and availability of all certification-related information and processes, including those involving AI systems and data.

When integrating AI into certification processes, programs must restrict access to AI systems and data to authorized personnel, implement continuous monitoring and logging of AI system activities, and conduct regular security audits to ensure ongoing compliance with security policies and standards. Data must be protected throughout its lifecycle.

Standard 13: Panel Composition

Certification programs must utilize panels of qualified SMEs for critical functions such as job analysis, item development, standard setting, scoring, and other examination-related activities. While AI technologies might support these processes, such as by analyzing panel compositions or summarizing discussions, they must not replace human judgment. Qualified human personnel must make final decisions regarding panel composition and related activities to ensure the integrity and validity of the certification program.

While AI can enhance efficiency in managing panel-related data and processes, it cannot replace the critical judgment and expertise of a SME panel or panel member.

Standard 14: Job Analysis

While AI technologies might be used to perform tasks such as reviewing and analyzing job-related data, they must not replace the essential input obtained directly from certificants or the intended target audience.

AI-generated insights are supplementary and must undergo validation by qualified human personnel. Moreover, the accuracy of AI-generated insights depends on the representativeness of the datasets on which they are trained. The expertise of SMEs is crucial to interpret AI findings, provide context, and ensure that the identified competencies align with real-world practices.

Programs must document the involvement of AI-related tools and methodologies and the procedures for human oversight.

Standard 15: Examination Specifications

AI systems may contribute to the preparation of examination specifications, such as summarizing job analysis results, but final exam weightings must be determined by qualified human personnel based on the job analysis findings. Documentation of AI's role in examination specification development must be available for review and subject to oversight.

Standard 16: Examination Development

Integrating AI into item development may enhance efficiency but requires human oversight. AI can assist with generating and organizing items, but SMEs must review all AI-generated content to ensure accuracy, fairness, and compliance with psychometric standards.

Programs must take reasonable precautions to ensure that AI-generated content does not infringe on others' intellectual property and is not inadvertently made publicly available. These precautions should be documented and implemented in a manner that maintains compliance with confidentiality policies outlined in Standard 12.

If an AI system is used to translate test content to another language, qualified human personnel (generally bilingual SMEs) must review and approve this work before it is used.

Certification programs must document AI's role, establish oversight protocols, and regularly assess AI-assisted content and psychometric characteristics, as they would with human-generated items.

Standard 17: Setting and Maintaining Passing Standards

AI might be used to perform tasks such as analyzing performance data to identify trends and suggesting passing standard adjustments, but SMEs must critically assess AI recommendations to align them with certification objectives.

AI must not replace the professional judgment of psychometricians or other qualified human personnel in setting and maintaining passing standards.

Standard 18: Examination Administration

Programs utilizing AI to assist in exam proctoring, candidate monitoring, or administration functions must ensure compliance with requirements related to bias mitigation and candidate fairness. Human oversight remains essential and human proctors are required.

While AI may be used to help detect irregularities, it must not autonomously terminate or invalidate exams. Qualified personnel should oversee AI-assisted proctoring, addressing potential anomalies like false positives or technical issues.

Standard 19: Scoring and Score Reporting

AI might be used to support scoring and evaluation processes when its role is clearly defined, appropriately validated, and does not replace psychometrically sound examination methods. AI-driven scoring contributions must be transparent and verifiable. Programs must disclose AI's role in scoring to the public and ensure that pass/fail classifications are based on validated examination performance data rather than AI-determined thresholds.

While AI might be used to generate preliminary scoring outcomes, final pass/fail decisions must be subject to audits and validation by qualified human personnel to ensure fairness and accuracy. This oversight should include systematic reviews of AI-driven scoring processes, bias detection, and anomaly resolution rather than requiring individual human review of every score.

Standard 20: Evaluation of Items and Examinations

If AI is used to evaluate item and exam performance, it may be used to generate recommendations, but final item retention, revision, or removal decisions must be made by qualified human personnel. AI-assisted evaluation must be justified, transparent, and documented.

Standard 21: Maintenance of Certification

If AI is used in recertification processes, such as evaluating self-assessment tools, portfolio reviews, or continuing education activities, programs must document how AI contributions align with competency-based certification requirements. AI-driven assessments must meet the same quality criteria as traditional evaluations and undergo human validation to confirm accuracy, fairness, and relevance.

Standard 22: Quality Assurance

AI may be used to assist in quality assurance activities, with human oversight. For example, it might be used to detect biases, confirm reliability, and uphold ethical decision-making. Programs should use diverse datasets, conduct regular bias assessments, and implement mitigation strategies to prevent discriminatory outcomes as appropriate.

Programs must maintain clear documentation detailing AI system processes and quality assurance protocols.

Standard 23: Maintaining Accreditation

AI use by certification programs must be disclosed to the NCCA through initial applications, five-year reaccreditation applications, and annual reports. However, the integration of AI alone, other than in connection with essential certification functions, is not inherently considered a material change.

The use of AI systems is considered a "material change" if it results in modification to a fundamental aspect of the certification program, such as:

- the legal status or governance structure of the certification organization;
- the purpose, scope, or activities of the certification program;
- the purpose, scope, or objectives of any certification examinations;
- the program name and/or designation; and
- examination development, administration, and/or evaluation procedures[4]. (See, Standard 23, Commentary 2).

References:

NCCA Standards (2021)

[4] The use of AI in item writing or exam administration alone does not automatically require a notice of material change. However, if an AI implementation results in broader procedural modifications to examination development, such as changes to review processes, SME involvement, or validation methods, it may necessitate a notice of material change. Programs should assess whether AI involvement impacts established examination development and/or delivery procedures and, if so, submit a material change notice accordingly.