INDEPENDENT
CERTIFICATION
PROTOCOLS FOR

AI-Enabled Hiring & Recruiting Technologies
I. BACKGROUND
The Independent Certification Protocols for AI-Enabled Hiring and Recruiting Technologies ("Certification Protocols") supplement the Principles for Trustworthy AI in Recruiting and Hiring ("Principles") established by BBB National Programs and the Center for Industry Self-Regulation.

Capitalized terms in this document have the same meaning as in the Principles.

Participating Employers and Vendors may self-certify their compliance with the Principles through public notices and disclosures. Additionally, in order to enhance workforce trust, reduce risk, and demonstrate accountability and transparency, Employers and Vendors may also undergo an independent certification through a third-party certifying organization ("Certifier"). Certified organizations may display an approved trust mark or seal provided by the Certifier. For such independent certification to be recognized, the Certifier and its procedures must meet or exceed the standards specified in these Certification Protocols.

To avoid confusion, the Certification Protocols refer to the organization seeking certification as the “Applicant Business.” Such organization may serve the role of Employer, Vendor, or both under the Principles. For purposes of certification, the organization’s role vis-à-vis AI-powered hiring tools and individual applicants for employment will be considered.

II. GOALS OF INDEPENDENT CERTIFICATION
A. Promotes accountability by Employers and Vendors.
B. Provides flexible response to technology changes.
C. Protects intellectual property via confidential independent third-party review.
D. Fosters process transparency and coalesces best practices.
E. A trusted certification mechanism should:
   1. Require independent reviews of policies, procedures, and technical documentation;
   2. Ensure transparency via attestations from certified businesses in the notice provided at time of data collection, including explicit commitment to the Certification Protocols and Principles for Trustworthy AI;
   3. Clarify responsibilities between Employers and Vendors based on enforceable commitments to established best practices;
   4. Support rigorous algorithm testing using representative data sets;
   5. Standardize understanding of factors that should trigger enhanced auditing;
   6. Normalize expectations for routine certification and risk-based approach to assessments and audits; and
   7. Engage in ongoing monitoring and/or redress for complaints.

1 Independent certification reduces third-party risk and increases the feasibility of risk measurement when AI-enabled tools are used across organizations. This mitigates many challenges identified in NIST’s AI Risk Management Framework 1.0: “Risk can emerge both from third-party data, software of hardware itself and how it is used. Risk metrics or methodologies used by the organization developing the AI system may not align with the risk metrics or methodologies used by the organization deploying or operating the system. Also, the organization developing the AI system may not be transparent about the risk metrics or methodologies it used. Risk measurement and management can be complicated by how customers use or integrate third-party data or systems into AI products or services, particularly without sufficient internal governance structures and technical safeguards.” National Institute of Standards and Technology Risk Management Framework 1.0 (Jan. 2023), https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf at 5.
2 For examples of recommendations for AI bias audits, see the New York City Automated Employment Decision Tool Law (Local Law 144) and the White House Blueprint for an AI Bill of Rights, https://www.whitehouse.gov/ostp/ai-bill-of-rights/.
III. CERTIFIER REQUIREMENTS

A. Transparency
On its public website, the Certifier must include information regarding the Principles and the services it performs under the Certification Protocols.

This information must include:
1. information on or a link to the Principles and Certification Protocols;
2. detailed certification standards against which Applicant Businesses will be assessed, consistent with the Certification Protocols;
3. procedures for certification review and annual renewal, documenting how the Certifier meets the requirements outlined herein, including the possibility of denying and revoking certifications; and
4. a list of all currently certified organizations along with, for each, a description of the scope the certification and the most recent review date.

B. Impartiality
The Certifier must be free of actual or potential conflicts of interest. This means the ability of the Certifier to perform all tasks related to an Employer or Vendor certification and ongoing participation in the Principles free from influences that would compromise the Certifier’s professional judgment, objectivity, and integrity.

To meet this requirement, the Certifier should have internal structural and procedural safeguards, such as:
1. Written policies for internal review of potential conflicts of interest with Applicant Businesses and certified organizations.
2. Written policies for disclosure of potential conflicts of interest and, where appropriate, withdrawal of the Certifier from particular engagements. Such withdrawal should be mandatory in cases where the Certifier is related to the Applicant Business or certified organization to the extent that it would give rise to a risk that the Certifier’s professional judgment, integrity, or objectivity could be influenced by the relationship. At no time may a Certifier have a direct or indirect affiliation with any Applicant Business or certified organization that would prejudice the ability of the Accountability agent to render a fair decision with respect to their certification and ongoing participation.
3. Written policies governing the separation of personnel handling certification functions from personnel handling sales and consulting functions.

C. Expertise & Substantive Review
The Certifier evaluates the written policies, procedures, and technical documentation of Applicant Businesses against a set of published program requirements that encompass the Principles. The scope of the review is focused on those documents necessary to validate processes for testing AI Applications for consistency with the Principles.

To ensure a fulsome review, the Certifier must:
1. Possess appropriate expertise to review and verify policies, procedures, and technical documentation against program requirements.

---

3 The Certifier may be engaged to perform consulting or technical services for an Employer or Vendor other than services relating to their certification and on-going participation in the Principles.

The Center for Industry Self-Regulation (CISR) is a 501(c)3 non-profit organization (FEIN: 84-5021924) and is a DBA of BBB National Programs Charitable Foundation. 1676 International Drive, Suite 550 | McLean, VA 22102.
2. Engage in a comprehensive dialogue with the Applicant Business that includes, at a minimum, the production and review of all relevant policies, procedures, and technical documentation about AI Applications within scope of the certification, with an opportunity to remedy any gaps identified.

3. Any material statements and findings from the review should be signed by a senior officer of the organization with authority to bind it.

D. Confidentiality and Data Security

The Certifier must establish comprehensive written procedures designed to ensure the confidentiality of all information received from Applicant Businesses and certified organizations, including communications and documentation. The Certifier must also establish appropriate policies to ensure the security of all documents provided during the certification process, including a data retention policy.

E. Accountability

The Certifier must establish comprehensive written procedures designed to ensure the integrity of the certification process and to monitor the certified organization throughout the certification period to ensure compliance with the Principles. Such procedures should include, at minimum:

1. Mechanisms to enforce Certifier program requirements, whether by contract or other means, including the ability for the Certifier to revoke the certification at any time due to unremedied findings of non-compliance.

2. Ongoing monitoring of transparency and disclosure requirements, such as verifying that public notices include required information.

3. A mechanism for receiving and resolving individual complaints related to the certified organization’s conduct under the Principles. Such mechanism must be monitored by personnel separate from those who conduct certifications.

4. Where there are reasonable grounds for the Certifier to believe that a certified organization has engaged in a practice that may constitute a breach of the Principles or Certification Protocols, an immediate review process will be triggered whereby verification of compliance will be carried out. Where non-compliance with any of the program requirements is found, the Certifier will notify the organization outlining the corrections it must make and a reasonable timeframe within which the corrections must be completed.

IV. CERTIFICATION PROTOCOL

A. Outline of Certification Steps

Stage 1 – Production of the application file, including organization self-assessment and supporting documentation such as written policies, procedures, and technical documentation of AI Application.

Stage 2 – Initial review of file against certification requirements.

Stage 3 – Feedback on non-conformity or anomalies, if any, and request to seek additional independent audit of AI Application, if necessary, based on risk.

Stage 4 – Secondary review of outputs/findings after opportunity to remedy identified gaps.

Stage 5 – Issue/deny certification and communicate results to organization in a comprehensive report, along with any other documentation needed to verify certified status. If approved, issue certification seal, verify public transparency and notice requirements are met, and list certified organization in public directory.

Stage 7 – Conduct annual re-certification reviews.
B. Application File

The application file submitted by the Applicant Business and reviewed by the Certifier must include the following elements:

I. Design process documentation:

Applicant Business shall document the design, development, and testing protocols for all AI Applications within the scope of the certification. This documentation shall include internal policies, procedures, and records of assessment sufficient to demonstrate the requirements of these Certification Protocols. If the Applicant Business does not have oversight or control over a listed criterion, it may rely on the separate certification of the Vendor that controls that aspect of the design, development, and testing of the AI Application. Alternatively, it may provide documentation of industry-standard internal validation processes, consistent with Section V.D.2. of the Principles.

Provided documentation will specify, in detail, the following elements:

a. Assessment of inputs:
   (1) types, formats (images in png or jpg format, audio mpeg or mp3, specific structures in xml, json, csv, or specific binary format, etc.) and the compatibility of these formats with other solutions or environments (in the case where interoperability is an issue),
   (2) source/acquisition method (sensors embedded into the hardware solution, data imported from a database, etc.) used by the AI functionality (if the source depends on the use cases, the range should be specified),
   (3) volume and quality of training data sufficient for the performance requirements and risks associated with the use of the AI functionality,
   (4) type(s) of algorithm(s) as well as the learning paradigm used by the AI functionality must be documented with respect to performance, maintenance and explainability constraints, frequency and output flow (the specifications of the learning paradigm (especially in the case where re-learning is possible after deployment),
   (5) the prototype of AI functionality whose normal operating domain of use and performance will be evaluated,
   (6) end-user requirements and expectations for AI functionality,
   (7) applicable regulatory requirements and how the AI Application meets these requirements,
   (8) complete description of how internal processes and procedures for reviewing inputs are consistent with requirements from similar activities and/or uses of AI Applications (standards, best practices etc.), and

---

4 Documentation can be specific, in the event the AI application is developed for a particular end-user, or generic, in the event it is expected to be widely distributed. A model can be built or adjusted based on data processed either manually by humans or using automated tools like machine learning algorithms, or both. Model Building often uses Historical Data/ Memory to aggregate data automatically into the Model but can also use Expert Knowledge. Objectives (e.g. the output variables) and Performance Measures (e.g. accuracy, resources for training, representativeness of the dataset) guide the building process. OECD Expert Group on Artificial Intelligence, Nov. 2019

5 For a list of factors that written policies and procedures should address, see National Institute of Standards and Technology Special Publication 1270, Towards a Standard for Identifying and Managing Bias in Artificial Intelligence (March 2022), https://nvlpubs.nist.gov/NISTpubs/SpecialPublications/NIST.SP.1270.pdf at 43 “Policies and Procedures.”

6 See NIST Special Publication 1270 at 17 “Dataset Guidance.”

7 (Title VII of the Civil Rights Act of 1964 (Title VII), the Americans with Disabilities Act of 1990 (ADA), and the Age Discrimination in Employment Act of 1967 (ADEA) and other applicable regulations prohibit the use of discriminatory employment tests and selection procedures.
(9) the presence of sensitive data, with discriminatory potential, should be checked and appropriate safeguards should be documented.

b. Assessment of functions:

(1) Data quality. The collection and use of training, validation, and test data must comply with applicable regulations. For example, in the United States, compliance with the Uniform Guidelines on Employee Selection Procedures, 29 C.F.R. Part 1607.

(2) Segmentation. The data used for testing must be distinct from the data used for learning (training and validation/development). The segmentation methods (random, by date, by configuration, or based on objective criteria (date, age, etc.) or randomly etc.) of the training, validation, and test datasets must be documented.

(3) Training and test data must be unique and not poisoned. The methods used to ensure that the databases are not poisoned and do not contain duplicates must be documented.

(4) In the case of learning on structured data, an effective method of handling (detecting and dealing with) missing values among critical attributes in the training and test data shall be implemented and documented.

(5) The learning database must include a training set and a validation set. Both sets must individually ensure a sufficient coverage of the system's use cases with respect to the performance requirements and risks associated with the use of the AI functionality, limit selection biases, and be balanced with respect to the use cases and objects under study.

(6) The learning database shall incorporate rare events or, alternatively, the AI functionality shall be capable of detecting the occurrence of such events. The Vendor shall document the list of rare events incorporated into the database along with their frequency of occurrence and the method used to determine them.

(7) The test database shall ensure sufficient coverage of the system use cases with respect to the performance requirements and risks associated with the use of the AI functionality.

(8) The test database shall include rare events. The Vendor shall document the list of rare events included in the database as well as their frequency of occurrence and the method used to determine them.

(9) All data transformation steps must be documented. If not, a study must be performed to evaluate the bias introduced by the pre-processing of the test database and the input data of the AI functionality of the deployed system and/or to reduce the impact of this bias. For example, the data normalization process is an example of pre-processing to be listed.

(10) If the learning process can generate biases, debiasing algorithms in the pre-processing, processing, or post-processing phase must be used, and this use must be documented and the records retained in accordance with applicable record retention schedules.

(11) An evaluation protocol based on performance measures must be implemented during and at the end of the learning process and documented (including human intervention for performance checks). 8

(12) Verification of compliance with the regulatory requirements established during the design process shall ensure that the AI functionality is not validated if any of the requirements are not met.

8 NIST recommends a holistic test, evaluation, validation, and verification (TEVV) approach. See generally NIST AI RMF 1.0 and NIST Special Publication 1270.
c. Assessment of outputs:

(1) Specifications of the functionality for which the AI Application is designed and tested, documenting:

(a) the specifications of the system (e.g., for a resume filtering process, the description could be: “the AI Application aims to identify a range of candidates for a position based on a matched review of the required criteria with the data supplied on the resume”),

(b) the use of the AI functionality (the automated task and its purpose), for example, “the required position criteria were met or not met”,

(c) levels of autonomy (human actions and controls on automated tasks),

(d) levels of performance where appropriate, in terms of: accuracy, reliability, learning execution time / execution time of the automated task on target hardware, resilience, including to attacks and outliers,

(e) and other examples of macro tasks (classification, ranking, regression, etc.);

(2) Requirements for associated documentation (for example, user documentation must be defined. It can also be defined that the test plans for the verification of the requirements concerning the AI functionality be included at specified intervals);

(3) Justification of the acceptability criteria for each of the requirements defined;

(4) Evaluation protocols, tools and metrics, the results of the evaluation process, the AI functionality whose field of use and performance will have been validated, and final risk analysis. The results of the evaluations (and any deviation from the requirements defined during the design process) must be documented; and

(5) Detailed description of risk mitigation efforts used in prior testing to avoid or suppress biased outputs and detail follow-on process testing schedules to ensure mitigation efforts maintain rigor.

2. End-user (Employer) documentation

If the Applicant Business plans to serve as a Vendor under the Principles, offering covered AI Applications for use by other organizations, it shall provide to Certifier detailed documentation for communication with the Employer in standardized format, including:

a. A description of the desired end-use of the AI Application along with instructions on its proper/stated use for end-users.

b. Appropriate information about testing procedures deployed (for example regarding the origin and composition of the training databases, the results of the evaluations, certainty levels of the functionality, its open- or closed-source character, and the specifications of the input data).
c. Descriptions of how the Principles have been met for the AI Application, including the following elements, as applicable:

(1) Confidentiality and respect for privacy, and access to data,

(2) Levels of transparency, in terms of:
   (a) Explainability and interpretability (including what is required and how long data should be kept),
   (b) Traceability, auditability of learning, and/or results,
   (c) Non-discrimination, and bias mitigation,
   (d) Results of evaluations,
   (e) Regulatory constraints,
   (f) Specifications related to the potential consequences of a failure of the AI functionality.

(3) Security of the AI Application.

d. Templates for documenting and communicating with the Employer about the AI Application, to include all of the following:
   (a) Clear indicators for validation.
   (b) Description of the post-processing (suggested or built-in) for the use of the AI Application.
   (c) Resources required and the operations to carry out calibration and maintenance (recommended or mandatory) after deployment.
   (d) The availability of data needed for learning to achieve the targeted performance and confidence levels validated or mentioned as a necessary condition for using the AI Application.
   (e) Contraindications (for example, if tests showed that the algorithm did not work for people with hyphenated surnames, then hyphenated surnames are a contraindication).

3. Risk analysis

The Applicant Business shall provide Certifier with a copy of the risk analyses performed at the pre-design, design, development, and deployment phases of covered AI Applications, as relevant and adapted to the functions of the AI Application. The risk analysis must identify, evaluate, and document the risks associated with the use of the AI Application and their potential impacts. This analysis must foresee the use of erroneous data that may be due to formatting errors, bugs in the data management system, or cyberattacks, and cover the components and sub-components as well as the interfaces between components of the AI functionality. The different failure modes of the AI functionality and their consequences must be established to allow the customers to be aware of the residual risks they are exposed to. The impacts can be quantified in terms of cost, safety, security, discrimination, etc.

C. Ongoing Commitments:

To receive and maintain certification, the Applicant Business and certified organization shall commit to the following practices:

1. To make available the specifications of functionality to anyone involved in the design, development, evaluation, or maintenance of the AI Application. For example, in the case of a Vendor, the specifications of the AI Application can be made available to Employers through the product sheet, available on the Vendor’s website, in the customer documentation, etc.

2. To document the design assumptions made about the AI functionality (including statistical assumptions that may vary over time) and the approach taken to select model types and evaluate the AI functionality.
In the event of a change in the assumptions or requirements applicable to the AI functionality, the certified organization shall ensure that an impact analysis is performed to ensure that the changes do not negatively impact compliance with the requirements of the Principles or the Certification Protocols.

3. To conduct outside audits of AI Applications, if recommended by Certifier based on the identified risks, with results reported during review.

4. To notify Certifier and provide fulsome documentation, if changes negatively impact compliance or otherwise would affect the disclosures and commitments made during certification.

5. To respond to individual inquiries or assist the Employer in responding to inquiries, as relevant, and cooperate with any dispute resolution procedures established by Certifier.

6. To maintain active certification by pursuing annual review and re-certification from Certifier.

---

11 The Certification Protocols embrace a risk-based approach to the review of AI Applications. Independent audits always serve to increase accountability and transparency. However, they are not a necessary precondition for certification except in situations where the Certifier identifies a high level of risk, consistent with best practices and applicable regulations.
About the Center for Industry Self-Regulation

The Center for Industry Self-Regulation (CISR), BBB National Programs’ 501(c)(3) nonprofit foundation, was created to harness the power of independent, industry self-regulation to address the marketplace trust challenges businesses face today. CISR is dedicated to education and research that supports responsible business leaders developing fair, future-proof best practices, and to the education of the general public on the conditions necessary for industry self-regulation.

About BBB National Programs

BBB National Programs, a non-profit organization, is the home of U.S. independent industry self-regulation, currently operating more than a dozen globally recognized programs that have been helping enhance consumer trust in business for more than 50 years. These programs provide third-party accountability and dispute resolution services that address existing and emerging industry issues, create a fairer playing field for businesses, and a better experience for consumers. BBB National Programs continues to evolve its work and grow its impact by providing business guidance and fostering best practices in arenas such as advertising, child-and-teen-directed marketing, data privacy, dispute resolution, automobile warranty, technology, and emerging areas. To learn more, visit bbbprograms.org.