CHAI Assurance Service Provider Certification Handbook v.1

2025 Edition

Brenton Hill

CHAI Assurance Service Provider and the CHAI logo

The term Assurance Service Provider and the CHAI logo are registered marks of the Coalition for Health AI, inc., which retains exclusive rights to control the use thereof. Permission to use the term and symbol is granted to CHAI-certified laboratories for the limited purpose of announcing their certified status, and for use on reports that describe only testing and calibration within the scope of accreditation. CHAI reserves the right to control the quality of the use of the CHAI Assurance Service Provider term, logo, and symbol.

For inquiries, contact CHAI:
By mail at PO BOX 809 Sudbury MA, 01776, or by email at admin@chai.org.

1 FOREWORD

1.1 CHAI ASSURANCE SERVICE PROVIDER CERTIFICATION PROCESS

Assurance Resources help accelerate the market adoption of AI models by establishing processes and tools developers can use to aid in streamlining and accelerating the adoption of AI in healthcare. CHAI is tasked with developing an evaluation framework for AI solutions in healthcare using consensus-driven standards and best practices. To help developers to trust these assurance resources, CHAI is certifying assurance resource providers who meet the following standardized criteria in the following areas:

1. Disclosing and minimizing conflicts of interest;
2. Protecting developer's intellectual property;
3. Ensuring standardized disclosures and processes for data controllers/processors;
4. Demonstrating that the data provided to the model developer meets the data specification /requirement required by the model developer;
5. Ensuring that the model's performance report meets the defined parameters approved by the model developer; and
6. Optional publishing of objective performance metrics using a CHAI model card on the CHAI National Registry.

As such, these resource providers will play a critical role in establishing trust in AI technologies by providing transparency of model performance within specific environments and by providing AI users and beneficiaries with more information about how this technology was evaluated and how it performed in other environments. These Assurance Resources are not part of any government regulatory process and are purely led by the private sector.

1.2 FRAMEWORK COMPONENTS

1.2.1 **Conflict of Interest disclosure.**

All CHAI-certified Assurance Service Providers sign a conflict-of-interest disclosure clause as part of the contracting process to be "CHAI-certified." This term will be required to flow down into the service agreements between the Assurance Service Provider and the Customer of the data.

1.2.2 **Privacy and Security of IP and Data.**

All CHAI-Certified Assurance Service Providers will be required to have privacy and security measures in place to safeguard the data and other confidential information used and produced in the assurance process. To ensure this happens, all Assurance Service Providers will be required to have, at a minimum, a SOC2 Type II Certification. Those with a

SOC2 Type I and that are in the process of working towards a Type II may also be considered for certification.

1.2.3 **Data Quality and Understanding.**

Data quality and integrity are necessary components to demonstrate robust and generalizable AI/ML models.

- Data attestation. All CHAI-Certified Assurance Service Providers shall ensure that the following are disclosed and included in the Model Card (if applicable): 1) data sources and providers, and 2) data attestation from each data provider that the data meets the Customer's Data Specification requirements.
- Data catalog. If possible, the Assurance Service Provider shall provide a data catalog for Customers to understand the types and attributes of data available. The data should be in a table format in alignment with USCDI V3 (see Figures 1&2). If data in the catalog is preformed, the data catalog should include attestations on data quality and integrity (See Figure 3).

1.2.4 **Use of CHAI Testing Frameworks.**

For CHAI-Certified Assurance Service Providers that offer validation reports, each will be required to use CHAI's Use Case-Specific Testing and Evaluation Frameworks and include these metrics and methodologies in the ensuing validation report.[1]

2 GENERAL INFORMATION

2.1 PROGRAM DESCRIPTION

The CHAI Assurance Service Provider Certification addresses the growing need for attested validation of AI/ML solutions and systems in healthcare. The program accredits organizations and resources capable of 1) enabling an infrastructure to securely utilize data for testing and conducting performance and conformance testing of AI/ML models as specified by the Customer, 2) maintaining an auditable record of the test artifacts (e.g., algorithm, data sets, performance report outputs), 3) assisting the Customer in building the Model Card based on the artifacts included in the enabling infrastructure, and 4) attesting to the accuracy and validity of the Model Card elements derived from artifacts collected and stored in the enabling infrastructure.

CHAI reserves the right to expand its certification program to include additional testing requirements or methodologies as industry standards evolve. Assurance Service Providers will be kept abreast of CHAI's updates on new frameworks and program offerings.

Certification under this program is a prerequisite for Assurance Service Providers aiming to provide CHAI-endorsed assurance services. However, achieving CHAI Certification does not authorize

[1] See Sepsis Risk Prediction T&E Framework for example.

immediate operation under specific regulatory programs (e.g., premarket clearance for a medical device) unless additional authorizations or approvals are obtained from the relevant regulatory bodies.

Maintaining CHAI Certification requires adherence to ongoing quality and ethical standards as outlined in this document and other applicable CHAI policies. Noncompliance may result in suspension or revocation of certification status.

2.2 DEFINITIONS

Assurance Resource. A tool, software program, or platform ("Resource") that enables Customers to validate performance of AI more effectively. Examples include enabling data infrastructure components, training and validation services, governance platforms, etc.

Assurance Service Provider. An organization ("Provider") that provides and maintains an Assurance Resource, and that uses that Resource and other assets to provide the assurance outputs defined by CHAI in this document.

Customer. Any individual or organization that engages the services of a CHAI-certified resource for AI assurance activities.

Certification. Formal recognition that a Provider is competent to perform relevant assurance functions using its Resource, as it pertains to assessments related to AI/ML systems in healthcare. Each Provider will be required to clearly state its intended use or value proposition.

Competence. The ability of a Provider to conduct its functions in accordance with specified standards, producing accurate and replicable results that meet intended purposes.

3 CONFLICT OF INTEREST POLICY AND DISCLOSURES

3.1 2.1 BACKGROUND

Each CHAI-Certified Assurance Service Provider will be required to adopt the Conflict of Interest Policy as articulated below and as referenced in the Assurance Service Provider contract.

3.2 2.2 POLICY

3.2.1 Purpose

This Conflict of Interest (COI) Policy establishes guidelines to ensure transparency, integrity, and ethical handling of data provided for the training and validation of artificial intelligence (AI) models in healthcare. The policy aims to prevent conflicts of interest that could compromise data quality, AI model reliability, or patient safety.

3.2.2 Scope

This policy applies to all entities, including healthcare providers, research institutions, technology companies, and data vendors involved in supplying, curating, or processing healthcare data for AI model development.

3.2.3 **Definitions**

- Conflict of Interest (COI): A situation where personal, financial, or professional interests may compromise or appear to compromise the integrity of data provision and AI model training and/or validation.
- Covered Individuals/Entities: Organizations and individuals engaged in the collection, sharing, or management of data used in AI healthcare applications.
- Sensitive Data: Any patient or proprietary healthcare information that must be handled according to legal, ethical, and regulatory standards.

3.2.4 **Disclosure Requirements**

- All entities providing data must disclose any financial interests, partnerships, or affiliations that could create a conflict.
- Researchers and employees involved in AI model development must report any relationships with data providers that could bias model outcomes.
- Any potential conflicts must be declared in writing before entering into service level agreements.

3.2.5 **Data Integrity and Objectivity**

- Entities must ensure that data is provided without manipulation, bias, or selective reporting to influence AI model outcomes.
- No entity should benefit financially or competitively from intentionally providing misleading or incomplete data.

3.2.6 **Independence of AI Model Validation**

- Assurance Service Provider attests that data sources and AI model training methodologies remain free from conflicts of interest unless disclosed.
- Entities supplying data should not have exclusive control over the validation and testing process.
- Any entity involved in both data provision and AI model commercialization must implement firewalls to ensure separation of interests.

3.2.7 **Prohibited Practices**

- Withholding adverse data that could impact AI performance.
- Financial incentives tied to specific AI model outcomes.
- Undisclosed partnerships between AI model developers and data suppliers.
- Use of proprietary or patient data without explicit consent and compliance with regulatory requirements.

3.2.8 **Reporting and Enforcement**

- Any suspected conflicts must be reported to CHAI immediately upon recognition.

- Violations of this policy may result in loss of CHAI-Certified status.
- Regular compliance audits will be conducted to ensure adherence to this policy.

3.2.9 **Acknowledgment and Compliance**

- All entities must acknowledge and agree to this policy as a condition of providing CHAI-certified Assurance Service Providers in each provider's service level agreements.
- Training and updates on COI policies will be provided to all stakeholders to ensure ongoing compliance.

3.3 Loss of Certification Status

Any Assurance Service Provider found to be in willful or negligent violation of this policy will be subject to the following corrective action plan:

1. **First Violation:** Resource provider will receive a written warning from CHAI with instructions on how to correct the mistake.
2. **Second Violation:** Resource provider will receive a written warning from CHAI with instructions on how to correct the mistake AND will be de-listed from the list of CHAI-Certified providers for 3 months with proper mitigation action plan.
3. **Third Violation:** Resource provider will receive a written letter notifying it of its immediate removal from the list of CHAI-Certified providers with a chance to re-certify in one years' time.

4 Code of Conduct

4.1 Introduction & Purpose

The purpose of this Code of Conduct is to maintain high standards of integrity, professionalism, and excellence within our CHAI-certified Assurance Service Providers. This section outlines the ethical and professional behavior expected of all employees, contractors, and stakeholders of Assurance Service Provider providers.

4.2 Ethical Standards & Core Values

- Integrity & Honesty
- Quality & Excellence
- Accountability
- Respect & Collaboration

4.3 Compliance with Regulations & Policies

- Follow all applicable local, national, and international laws (if pertinent) and regulations (e.g., HIPAA, GDPR, etc.).
- Remain in good standing with all pertinent regulatory bodies (HHS, EU, etc.)

- Use CHAI-approved testing methods and metrics (where provided)

4.4 Workplace Conduct & Safety

- Treat colleagues, clients, and partners with respect and professionalism.
- Ensure a discrimination- and harassment-free workplace.

4.5 Data Integrity & Reporting

- Provide supporting evidence that processes applied for assurance resources are reproducible and consistent.[2]
- Prohibit falsification, manipulation, or selective reporting of data.
- Maintain secure and traceable documentation for all Assurance Service Provider activities.
- Provide supporting evidence of use of CHAI Testing and Evaluation Frameworks where applicable.[3]

4.6 Conflicts of Interest

- Disclose any financial, professional, or personal relationships that may influence impartiality.
- Avoid situations where personal gain could affect Assurance Service Provider credibility or independence.

4.7 Use of Assurance Service Providers

- If the Assurance Service Provider wishes to market any net new, non CHAI-certified offerings, it must first work with CHAI to certify such service using the same processes found in this certification manual. Under no circumstances will use of CHAI's certification mark be permitted for non-certified offerings.
- Prevent waste, theft, or unauthorized use of resources.

4.8 Confidential Information

- CHAI and Providers will share confidential information with one another on an "as needs to know basis." This can include but is not limited to the information required for certification to occur or other pertinent information not currently contemplated.
- Providers will maintain all Customer information as Confidential Information, unless otherwise specified by agreement between the parties
- CHAI will not receive the confidential results or outputs generated as part of an assurance process and held by the Provider unless the Customer acknowledges in writing that they would like to share it for publication on CHAI's National Registry or other situations as agreed to in writing.

---

[2] The hallmark of a CHAI-Certified Assurance Resource Provider is ensuring customers anticipate receiving a consistent product each time they engage the provider. Providing attestation of this could mean including organizational policies on processes in place, leveraging a platform to provide the service, or other reasonable means. This will be considered on a case by case basis at the inception of the CHAI certification process.

[3] See Sepsis Risk Prediction T&E Framework for example.

4.9 Whistleblower Protection & Reporting Misconduct

- Assurance Resource Providers should have a whistleblower policy in effect and will provide supporting evidence to CHAI of it's application before certification is given.

4.10 Administrative Actions

- Violations of this Code may result in revocation of the Provider's CHAI-Certified status.

4.11 Acknowledgment & Compliance

- All CHAI-Certified Providers will acknowledge required adherence to this code of conduct in the Assurance Service Provider contract and in contracts with customers sourced through the CHAI Assurance Resource Marketplace.
- If the Provider is out of compliance with this code of conduct or has reason to believe it might be, it must contact CHAI leadership immediately.

5 Privacy and Security Certification

5.1 Background

Due to CHAI's commitment to ensuring that CHAI-Certified Assurance Service Providers are trustworthy and competent, each resource provider will be required to achieve a SOC II Type II designation. A SOC 2 Type II certification is crucial for organizations handling sensitive data, as it demonstrates adherence to security, availability, processing integrity, confidentiality, and privacy standards over time. It builds customer trust, reduces security risks, ensures regulatory compliance (e.g., HIPAA).

5.2 System Description

The system description in the SOC 2 Type II certification shall incorporate the Resource used in the Assurance Service.

5.3 Providing Proof of Certification

Each Assurance Service Provider will be required to submit an official copy of its SOC II Type II Certification as part of its initial onboarding as a CHAI-Certified resource provider.

5.4 Working Towards Certification

CHAI will exercise its certification discretion for providers that have a SOC II Type I certification and are working towards a Type II certification.

5.5 Loss of Certification Status

Any CHAI-Certified Assurance Service Provider that is no longer SOC II Type II Certified will notify CHAI immediately of its nonconformance with a corrective action plan on how it will regain certification and the timeframe for such plan.

Those that let the SOC II Type II Certification lapse, either willfully or negligently, will be subject to the following corrective action plan:

1. **First Violation:** Provider will receive a written warning from CHAI with instructions on how to submit a corrective action plan.
2. **Second Violation:** Provider will receive a written warning from CHAI with instructions on how to submit a corrective action plan AND will be de-listed from the list of CHAI-Certified providers for 3 months or until corrective action plan is put in place.
3. **Third Violation:** Provider will receive a written letter notifying it of its immediate removal from the list of CHAI-Certified providers with a chance to re-certify in one years' time.

6 Data Quality and Integrity

6.1 Background:

High quality data and processes to ensure transparency in data provenance and methods of data forming of this data is a hallmark high quality algorithm validation. If data is missing, not consistent, or is unreliable, the evaluations are less likely to be accurate or reliable when assuring for its intended purpose.

This section is intended for those Assurance Service Providers that maintain data sets with the intended purpose of providing it to customers for AI solution training and validation. The content for this section has been derived from FDA draft guidance and[4] AHRQ guidance,[5] which provide considerations for ensuring relevance and reliability of data when using real-world evidence to support regulatory decision-making for medical devices. CHAI is using this draft guidance to create a series of attestations the assurance providers will be attesting to during the initial phase of the certification.

Assurance providers that do not maintain data set but rather facilitate access to data held by data controllers, are encouraged to index the data that is available from the data controllers through data catalogs that meet this guidance.

The final responsibility for data set content and quality is jointly held by the data provider and the customer. The customer is ultimately responsible for the data specification and accepting the quality of the data as attested by the data provider. The customer is responsible for ensuring the data specification addresses the parameters listed in Figure 3 as they require to validate their product claims and indications.

6.2 Instructions

6.2.1 Assurance providers that maintain data sets

Each assurance provider that maintains data sets will be required to provide answers about data it makes available to customers:

1. Table 1 of data elements and attributes found in the data source (see Figure 1)
2. Included in the Table 1 should be a description of the data formats or types (see Figure 2)

[4] https://www.fda.gov/media/174819/download

[5] https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/registries-guide-3rd-edition_research.pdf

3. Attestations of data quality and provenance (see Figure 3)

    a. Relevance of Data: Datasets must reflect current clinical practices and use cases.

    b. Generalizability: Datasets should include diverse populations to reduce bias.

    c. Timeliness: Data must be regularly updated to align with evolving medical standards.

    d. Accrual Consistency: Collection methods must be standardized across all sources.

    e. Audit Trails: Traceability from data collection to analysis must be maintained.

    f. Bias and Fairness Checks: Regular audits to identify and mitigate bias in datasets.

    g. Security Compliance: Robust measures like encryption and access controls to protect data.

6.2.2 **Assurance providers that enlist data controllers**

Each assurance provider that enlists data controllers to provide test data shall ensure the data controller attests to the content and the quality of the data as specified by the customer in a form that meets model card format requirements.

7 ROLES AND RESPONSIBILITIES

7.1 CHAI

1. Publish the CHAI Assurance Provider Certification Handbook.
2. Publish the CHAI Code of Conduct.
3. Publish the CHAI Conflict of Interest Policy.
4. Certify Assurance Providers.
5. Review of the Assurance Provider outputs (model cards) before publishing
6. Create and maintain a public-facing and searchable database of AI model cards
7. Audit Assurance Provider activities and ongoing certification.

7.2 CUSTOMER

The Customer is any individual or organization engaging the services of a CHAI-certified Assurance Provider.

7.2.1 Customers engaging Assurance Providers for Data Processing, Training, and/or Validation

1. Internal: Develop the product under test, with internal controls that satisfy their intended customer base. At a minimum, the controls must include a method for product/model version control that allows traceability in the Assurance project.
2. Internal: Define the Clinical Data Plan and Statistical Plan (or equivalent) that satisfy their internal requirements, which may include:
    1. Claims and Intended Use Justification,
    2. Regulatory submission requirements,
    3. Proof of Generalizability within the Intended Population,
    4. Internal Quality Control requirements,
    5. Internal Governance and Legal requirements,
    6. End user requirements.
3. Define the product's Intended Use and associated claims, including inputs and outputs
4. Define a Project Opportunity Statement that includes the Assurance Project's scope.
5. Define the Data Specification for the product under test. The Customer is responsible for ensuring the Data Specification meets their internal requirements.

ⓘ For those providing data processing services, the Customer is responsible for defining a data set that meets their needs to generate proof of their claims' validity to the satisfaction of their customers.

6. Complete the CHAI Applied Model Card and determine if it would like to host on CHAI's National Registry through its vendor portal.
7. Define and implement (in a reporting function contained in the algorithm) the model assurance outputs that establish and ensure that the model meets the intended use and user requirements.

7.2.2 Customers engaging other types of Assurance Providers[6] (additional details coming soon)

1. Work with Assurance Provider to determine proper use of services or product.
2. Enable Assurance Provider with proper and appropriate access to organizational technology and other elements.

7.3 Data Controller[7]

1. Enable assurance provider needed functionality within their cloud or on-premise environment.

[6] I.e., governance platforms, etc.

[7] Depending on the platform and service offering, the "Data Controller" can be the same or separate entity from the Assurance Provider.

2. Review the Project Opportunity Statement promptly. Respond to the Customer with the ability and willingness to participate (e.g., requested data is available).
3. Provide a quotation for the data requested.
4. In the context of a project:
    1. Curate the requested data, to the specification's requirements.
    2. Attest to the Customer that the data meets their requirements.
    3. Make the data available to the project in a secure way.
    4. Perform requested tests according to CHAI Testing & Evaluation Frameworks, if any (if offering training and/or validation testing)
    5. Review and consent to controls provided by the assurance provider.
5. Maintain the data set in immutable storage for the time required by the customer for the project.

7.4 ASSURANCE PROVIDER

1. Conform to the CHAI Assurance Provider Certification Handbook requirements, as modified per this updated framework (see above).
2. Create and maintain a system (product, environment, process, etc.) that
    1. provides a service to the Customer for training, testing, and/or validating their algorithms,
    2. meets the data security and privacy needs of the Data Controller,
    3. provides auditable traceability of the links between the product under test, the data set, and the data attestation.
3. Produce or assist the Customer in producing the model card report.
4. Provide a letter of attestation for the model card report elements contained in the test

    1. algorithm version
    2. data set attestation
    3. algorithm outputs
5. Maintain auditable records for the duration required by CHAI and the customer.
6. Maintain the product under test as submitted in a Customer specific repository for the duration required by CHAI and the customer.

8 PUBLISHING RESULTS ON CHAI NATIONAL REGISTRY

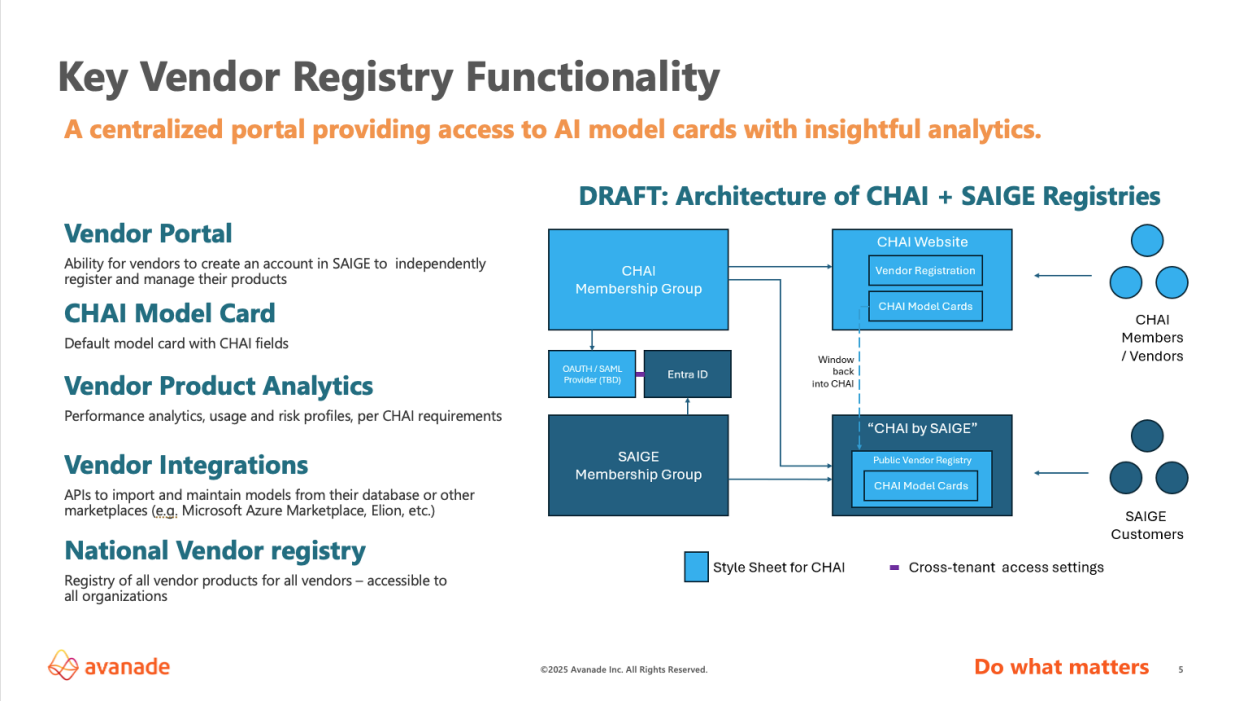
8.1 BACKGROUND

In order to provide the marketplace with an increased degree of trust in AI solutions, CHAI is developing a National Registry to provide CHAI-Applied Model Cards to the public, as an open-source resource. The registry will be optional for all developers and for all Assurance Resource Providers.

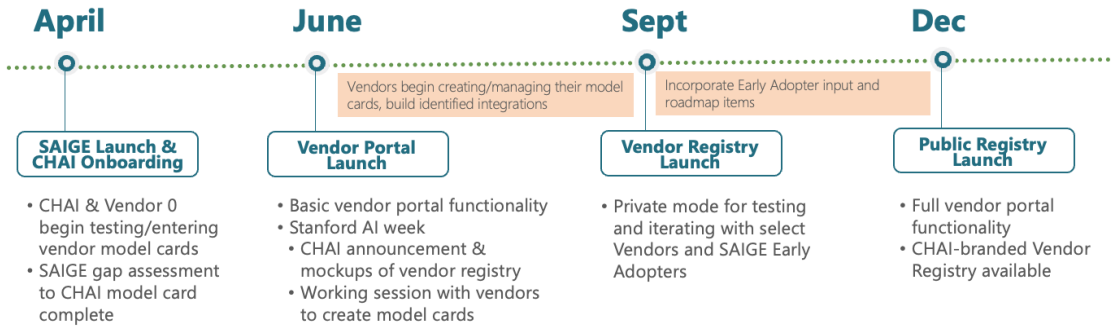
8.2 STEPS TO UPLOAD RESULTS TO THE NATIONAL REGISTRY

The CHAI National Registry is in the process of being built and the exact process for integrating Assurance Resource Provider outputs is still being determined.

The overarching desire and timeline are demonstrated in the graphics below:



Vendor registry timeline 2025



avanade CHAI

Confidential

Do what matters

6

9 ATTRIBUTES LISTED ON CHAI ASSURANCE RESOURCE WEBPAGE

9.1 BACKGROUND

In order to provide perspective customers with transparent and accurate information about each assurance resource provider, information found in section 8.2 will be required to disclosed on CHAI's public website. All disclosures will be reviewed before appearing on the CHAI website.

9.2 REQUIRED PUBLIC DISCLOSURES

- Primary Function of Assurance Resource
- Intended use of Assurance Resource (should include out of scope uses)
- Number of Current Customers
- AI Type Serviced (Traditional, ML, Generative, Agentic, etc.)
- Pricing Model
- Conflict of Interest Disclosures (in accordance with the policy in **Section 3**)
- Data Catalogue (in accordance with Section 6 and Appendix)[8]
- "Soc 2 Type II Certified" or "*Soc 2 Type I Certified working towards Type II*"

[8] The distinction as to whether the Data Controller and Assurance Provider are the same or different entities should be listed here per section 7.3 and 7.4.

10 APPENDIX

10.1 FIGURE 1

Data Class	Data Elements
Allergies and Intolerances	- Allergy/Intolerance Type- Allergy/Intolerance Substance (Medication)- Allergy/Intolerance Substance (Food)- Allergy/Intolerance Substance (Environmental)- Reaction- Severity
Assessment and Plan of Treatment	- Assessment and Plan of Treatment Note
Care Team Members	- Care Team Member Name- Care Team Member Identifier- Care Team Member Role- Care Team Member Location- Care Team Member Telecom
Clinical Notes	- Consultation Note- Discharge Summary Note- History & Physical- Imaging Narrative- Laboratory Report Narrative- Pathology Report Narrative- Procedure Note- Progress Note
Diagnostic Imaging	- Diagnostic Imaging Order- Diagnostic Imaging Report- Diagnostic Imaging Study
Encounter Information	- Encounter Type- Encounter Diagnosis- Encounter Time- Encounter Location- Encounter Disposition
Functional Status	- Functional Status- Disability Status
Goals	- Patient Goal
Health Concerns	- Health Concern
Health Insurance Information	- Coverage Status- Coverage Type- Relationship to Subscriber- Member Identifier- Subscriber Identifier- Group Number- Payer Identifier
Immunizations	- Immunization Name- Immunization Date- Immunization Manufacturer- Immunization Lot Number- Immunization Expiration Date- Immunization Site- Immunization Route- Immunization Dose Quantity- Immunization Status
Laboratory	- Laboratory Test- Laboratory Value/Result- Specimen Type- Result Status

Medications	- Medication Name- Medication Dose- Medication Dose Unit of Measure- Medication Route- Medication Frequency- Medication Indication- Medication Fill Status
Patient Demographics	- Patient Name- Patient Identifier- Patient Address- Patient Telecom- Patient Gender- Patient Birth Date- Patient Birth Place- Patient Race- Patient Ethnicity- Patient Preferred Language- Patient Occupation- Patient Industry
Patient Education	- Education Material- Education Material Date
Problems	- Problem- Problem Status- Problem Date of Onset- Problem Date of Resolution
Procedures	- Procedure- Procedure Status- Procedure Date- Procedure Performer- Procedure Device
Provenance	- Author Time Stamp- Author Organization- Author Role
Smoking Status	- Smoking Status
Social Determinants of Health	- SDOH Assessment- SDOH Goals- SDOH Interventions- SDOH Problems
Vital Signs	- Vital Sign Name- Vital Sign Result- Vital Sign Units- Vital Sign Date/Time

10.2 **FIGURE 2**

Technical and Data Requirements

CHAI-certified assurance providers that provide data sets directly work with a variety of data formats to ensure the validation of AI models across healthcare applications. These providers will be required to disclose data types and formats (in addition to the data elements found in Figure 1) to developers to ensure alignment with developer preferences and formats. See examples below:

1. **Structured Data**: Formats like CSV, HL7, and FHIR for clinical records and interoperability.
2. **Unstructured Data**: Formats like DICOM, PDFs, and plain text for imaging and notes.
3. **Semi-Structured Data**: Formats like JSON and XML for API-driven exchanges.
4. **Specialized Formats**: Including genomic data (VCF, FASTA) and imaging formats (TIFF, JPEG).

10.3 **FIGURE 3**

Instructions: Please review each requirement below and indicate compliance by checking the appropriate box. Supporting documentation should be provided where indicated.

Requirement	Description	Compliant (Yes/No)	Supporting Evidence (if applicable)
Relevance of Data	Degree to which the characteristics of the data satisfy stated and implied needs of the customer when used for a particular model type[9]	☐ Yes ☐ No	Provide details on the kinds of clinical and non-clinical use cases that you have relevant data for, and provide details of validation protocols to demonstrate relevance.
Generalizability	Datasets represent diverse populations to reduce bias and ensure inclusivity.	☐ Yes ☐ No	Include demographic breakdown of data.
Timeliness	Data is updated regularly to reflect changes in medical standards or clinical environments.	☐ Yes ☐ No	Submit update schedules for datasets.
Accrual Consistency	Data collection methods are consistent across all sources and sites.	☐ Yes ☐ No	Attach data collection protocols.
Audit Trail	Complete traceability from data collection to analysis is maintained.	☐ Yes ☐ No	Provide examples of audit logs.
Data Provenance	Documentation ensures clarity on data origins and any processing applied.	☐ Yes ☐ No	Submit metadata or lineage documents about the creation, update, transcription, abstraction, validation, and transferring control of any data.[10]
Bias and Fairness Checks	Regular evaluations are conducted to identify and mitigate bias in datasets.	☐ Yes ☐ No	Include results of bias audits or mitigation plans.
Security and Privacy Compliance	Robust measures (e.g., encryption, access control) ensure data	☐ Yes ☐ No	Provide details on security measures and compliance.

[9] See ISO 42001 section B.7.4

[10] See ISO 42001 section B.7.5.

	security and privacy compliance.		
Validation and Monitoring Procedures	Quality assurance measures are in place for data consistency, accuracy, and reliability.	☐ Yes ☐ No	Submit QA and validation reports.

Data Documentation Requirements

Instructions: Please review each requirement below and indicate compliance by checking the appropriate box. Supporting documentation should be provided where indicated.

Requirement	Description	Compliant (Yes/No)	Supporting Evidence (if applicable)
Data Collection Documentation	Records include purpose, scope, data sources, and collection processes.	☐ Yes ☐ No	Submit collection protocol or metadata.
Transformation Logs	Documentation of data cleaning, normalization, and any transformations applied.	☐ Yes ☐ No	Provide transformation workflow details.
Quality Control Checks	Documentation of accuracy, completeness, and consistency checks performed on data.	☐ Yes ☐ No	Include QC process and findings.
Version Control	All datasets and related documentation are version-controlled and traceable.	☐ Yes ☐ No	Submit version history for datasets.
Data Security Documentation	Evidence of data access logs, encryption protocols, and compliance with privacy standards (e.g., HIPAA/GDPR).	☐ Yes ☐ No	Provide security audit reports.
Provenance and Lineage	Detailed lineage tracking from raw data to processed outputs.	☐ Yes ☐ No	Include data lineage diagrams or reports.
Validation	Documentation of validation steps, including accuracy checks and model performance testing.	☐ Yes ☐ No	Submit validation protocols and results.
Audit Trail Maintenance	A clear, accessible audit trail is maintained for all changes to data.	☐ Yes ☐ No	Provide examples of maintained audit trails.
Reporting Documentation	Reports include assumptions,	☐ Yes ☐ No	Submit sample reports with required details.

	limitations, and statistical methods used in analysis.		