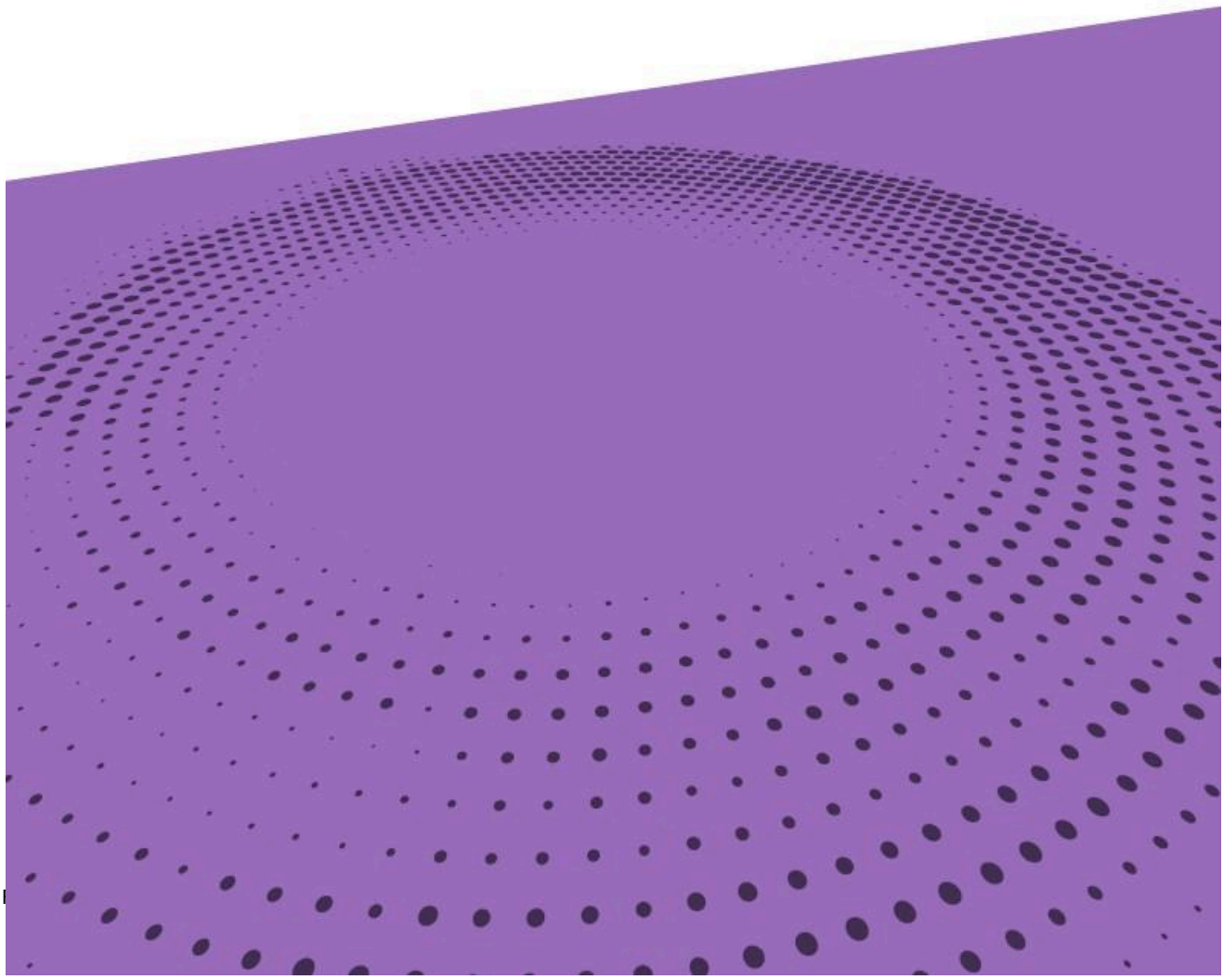
FORHUMANITY
980 Broadway #506 Thornwood, NY 10594
(+1) 9146028663
ryan@forhumanity.center
www.forhumanity.center

FORHUMANITY EUROPE
12 rue Frédéric Petit
80000 Amiens France

FORHUMANITY

EU AI Act – Provider only

CERTIFICATION SCHEME V1.5
Artificial Intelligence, Algorithmic and Autonomous (AAA) Systems



Introduction
Infrastructure of Trust
ForHumanity's Role in an Infrastructure of Trust
EU Artificial Intelligence Act
1.0 Scope
1.1.1 Determination of Provider Status
1.1.2 Jurisdiction(s)
1.1.3 Relevant Legal Frameworks
1.1.3.1 In scope - GDPR Applicability
1.1.3.2 In scope - European Accessibility Act of 2019 (EU) 2019/882
1.1.3.3 In scope - Digital Service (EU) 2022/2065
1.1.3.4 In scope - Cybersecurity minimum requirements
1.1.4 Conditions for using this certification scheme as an Importer
1.2 Audit Period of Validity
1.3 Out of Scope Systems
1.3 Target of Evaluation Determination Process
1.4 Territorial Scope
2.0 Normative References
3.0 Terms and Definitions
4.0 General Requirements for Accreditation
4.1 Interoperability with Standards
4.2 Normative Criteria explanation
4.3 Documentation of Assessments and Certification
4.4 Evaluation Methodology
5.0 Criteria catalogue
Expert Oversight
Top Management and Oversight Bodies
Relevant Legal Framework and Modular Assurance Assessments
Organisational Controls
Training and Education (AI Literacy)
Specialty Committees
Prohibited System - Article 5
Excluded AI Systems (Recital 53-60)
Business Rationale
General-Purpose AI Determination
Ethical Oversight
Consumer Protection

Excerpt-only

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

Data Privacy and Protection
Article 9 - Risk Management
Article 10 - Data Management and Governance
Bias Mitigation
Explainability
Technical Infrastructure
Design Choices
Accommodations - Recital 80
Choice Architecture
Security and Cybersecurity
Testing and Evaluation
Article 72 - Monitoring (Continuous and Post-Market)
Incident Management (Article 73.6)
Article 11 - Technical Documentation
Article 12 - Record Keeping - Logs
Article 13 - Transparency and Provision of Information to Deployers
Controls
Article 14 - Human Oversight and Interaction
Exceptions Interpretability
Article 15 - Accuracy, Robustness, and Cybersecurity
Vendor Management
Change Management
System Development Life Cycle
Article 17 - Quality Management System
Regulatory Compliance
Article 18 - Documentation Keeping
Decommissioning
Article 49 - EU Database Registration
Article 50 - Transparency Obligations
Article 51 - General-Purpose AI Classification
Article 52 - Notification of a General Purpose AI as Systemically risky
Article 53 - Obligations of a Provider of a General-Purpose AI Model
Article 55 - Obligations of a Provider of a General-Purpose AI Model with Systemic Risk
Article 60 - Real World and Beta Testing
Article 61 - Informed Consent for Real World Testing
Appendix A - Infrastructure of Trust for AI - Guide to Entity Roles and Responsibilities
Background on Independent Audit

Excerpt-only

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

Adapting to AI and Autonomous Systems
Role on Independent Audit of AI and Autonomous Systems
Participants in the System
Licensing

Excerpt-only

This document is the property of ForHumanity Inc. ©2022-24, ForHumanity Inc. a 501(c)(3) tax-exempt Public Charity

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

Introduction

ForHumanity (https://forhumanity.center/) is a 501(c)(3) non profit organization and ForHumanity Europe is a French 1901 Association, dedicated to addressing risks associated with Ethics, Bias, Privacy, Trust, and Cybersecurity in Artificial Intelligence, Algorithmic, and Autonomous (AAA) Systems. ForHumanity uses an open and transparent process that draws from a pool of over 2600+ international contributors to construct audit criteria, certification schemes, and educational programs for legal and compliance professionals, educators, auditors, designers, developers, and legislators to mitigate bias, enhance ethics, protect privacy, build trust, improve cybersecurity, and drive accountability & transparency in AAA Systems. ForHumanity works to make AAA Systems safe for all people and makes itself available to support government agencies and instrumentalities to manage risk associated with AAA Systems. Our mission is to *examine and analyze downside risk associated with the ubiquitous advance of AI, algorithmic and autonomous systems and where possible to engage in risk mitigation to maximize the benefits of these systems… ForHumanity*

Infrastructure of Trust

ForHumanity supports an infrastructure of trust predicated on the 50+ year track record of financial accounting and reporting. This infrastructure of trust is founded on a principle of jurisdictional sensitivity, which means that each sovereign nation-state or region has the right to establish their own laws, regulations, guidelines, and shared moral framework. ForHumanity affirms that right by ensuring that our certification program upholds local laws and seeks approval, where applicable, from local authorities. Key elements of Independent Audit of AI Systems are critical to ensure that it functions properly across multiple different jurisdictions, these are non-negotiable elements of the shared moral framework that constitutes Independent Audit of AI Systems and they include concepts such as transparency, disclosure, independence, risk management, and ethical oversight.

ForHumanity believes that a binary (compliant/non-compliant) set of criteria, either adopted by common practice in the marketplace or approved by the sufficient governmental authorities, and subsequently assured for compliance independently by certifying bodies (auditors), can create an infrastructure of trust for the public that assures compliance with laws, regulations, guidelines, standards, and best practices in a proactive manner when combined with the requirement for regular, mandatory, independent audits.

Excerpt Only

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

An infrastructure of trust, as it relates to certification, is an unconflicted process deploying a segregation of duties, conducted by certified and trained experts, that establishes a robust ecosystem that engenders trust for all citizens and protects those who have no power or control.

The infrastructure of Trust that For Humanity supports is grounded on four core tenets:

1. ForHumanity produces accessible, binary (compliant / not compliant) certification criteria that transparently and inclusively aligns laws, regulations, standards, guidance and best practice that embeds compliance and performance into practice, and is considerate of corporate wisdom, but impervious to corporate dilution and undue influence, while being mindful of the regulatory burden and dedicated to maximizing risk mitigations to humans.

2. Individuals are trained and accredited on certification criteria as experts by ForHumanity. They perform pre-audit and audit services on behalf of certification bodies and are individually held to a high standard of behavior and professionalism as described in the ForHumanity Code of Ethics and Professional Conduct - they are ForHumanity Certified Auditors (FHCAs)

3. Certification Bodies employ FHCAs to independently assure compliance with certification criteria on behalf of the public. They are licensed, independent, robust organizations that take on the task and risk, on behalf of the public, to ascertain assurance of compliance. They are held to standards of independence and anti-collusion and are further subject to third-party oversight ("watching the watchers"), by entities such as national accreditation bodies (e.g. COFRAC, UKAS, DaKKE) and ForHumanity.

4. Corporations and public sector Providers and Deployers of AAA Systems can use the criteria to operationalise governance, oversight, and accountability that helps them to achieve required conformity under the law. Compliance with ForHumanity certification schemes will create leverageable governance, oversight, and accountability that will simultaneously lead to more sustainable profitability and reduce the risk of negative outcomes for their stakeholders.

See Appendix A for more details on Roles and Responsibilities in an Infrastructure of Trust[1].

[1] Infrastructure of Trust for AI – Guide to Entity Roles and Responsibilities

Discretionary

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

ForHumanity's Role in an Infrastructure of Trust

Founded in 2016, ForHumanity first wrote about Independent Audit of AI Systems in 2017 and it has been our primary focus since that time. We advocate for mandatory independent audits and the establishment of the aforementioned infrastructure of trust similar to those required in financial accounts and reporting.

Transforming an audit ecosystem from financial audits to process audits for AAA Systems requires thoughtful adaptation. Transformation occurs by accomplishing the following tasks:

1. Understanding how financial audit rules & standards mitigate risk, provide clarity, and translate opaque controls and processes into public trust and valuable cross-sectional comparability through third-party independent assurance
2. Understanding the risks of AAA Systems and developing rules & standards to treat and mitigate risks to stakeholders, including individuals
3. Drafting audit criteria that are binary, implementable, solution-oriented to the identified risks
4. Mapping steps #1-3 onto an ecosystem that recreates the assurance and infrastructure of trust nurtured in financial audit for more than 50 years

In support of this transformation, ForHumanity is replicating and augmenting the role of the Financial Accounting Standards Board (FASB) and the International Financial Reporting Standards (IFRS) foundation, who drafted GAAP and IFRS respectively. Unlike those predecessors, ForHumanity is a grassroots, civil-society organization with contributors from more than 98 countries around the world. Our approach ensures globally-harmonized, audit criteria that operationalize the law, standards, and best practices sourced by diverse input and multi stakeholder feedback contributors.

We draft audit criteria for AAA Systems in the context of new legislation all around the world, such as, the EU's General Data Protection Regulation (GDPR), and the EU Artificial Intelligence Act, Consumer Protection and Consumer Privacy Protection Act in the United States, Lei Geral de Proteção de Dados Pessoais (LGPD) in Brazil, and India's Digital Personal Data Protection Act

ForHumanity's authority for producing audit criteria is grounded in the robustness of our crowdsourced, transparent process (no one is excluded from participating), however we always seek the endorsement of Federal, state, and local authorities, as applicable, when they support the approval of audit criteria, such as the manner in which most nation-states and regional blocks have adopted Generally Accepted Accounting Principles (GAAP) or International Financial Reporting Standards (IFRS) to govern financial accounting and reporting. When

Excerpt-only

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

governments are unprepared to endorse uniform, objective audit criteria, then ForHumanity seeks adoption directly from the marketplace, which is what occurred in 1973 with GAAP and the predecessor to IFRS, prior to Federal adoption in the years afterwards.

EU Artificial Intelligence Act

The EU Artificial Intelligence Act (EU AI ACT) - (EU) 2024/1689 is part of an intertwined set of laws and regulations that govern the use of artificial intelligence (AI) in the European Union[2]. This law prohibits certain uses of AI and regulates the use of AI that have been deemed to be a "high risk" to the health, safety, and fundamental rights of EU citizens. "High Risk" is defined in multiple places in the law, especially Annex I Section A & B and III and can be described in Recital 5 as being material or immaterial including physical, psychological, societal or economic harm. Further the law aims to govern the use of General-Purpose AI by both Providers and Deployers in the European Union, including a designation for General Purpose AI as systemically risky. The Act establishes a set of compliance requirements and regular oversight to ensure conformity. Conformity may be determined by self-assessment or third party assurance by notified bodies. Transparency and disclosure requirements are also included and must be submitted to national supervisory authorities.

AAA Systems are (often complex) socio-technical tools. As a result, the EU AI Act, and this certification scheme are considerate of other related laws and regulations. For example, any AAA System that is covered by the EU AI Act and uses Personal Data of Data Subjects in the EU, must also be concurrently compliant with the (EU) 2016/679 General Data Protection Regulation. Other laws, regulations, and guidance that impact the AAA Systems governed by the EU AI Act include:

1. General Data Protection Regulation
2. Digital Services Act
3. European Accessibility Act of 2019
4. EU Data Act

This interconnected web of laws and regulations establishes legal obligations groups known as, Providers, Deployers, Controllers, and Processors. This certification scheme is dedicated to the obligations of Providers. The ForHumanity EU AI Act Deployer-only v1.4 certification scheme is

[2] All references in this document to EU Member States, EU Data Subjects, the Union, Union Law, etc. shall be understood to include, in addition to their meaning in the regulation, European Economic Area (EEA) states, EEA Data Subjects, the EEA, and the EEA Agreement respectively

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

Excerpt only

to be used by Deployers of AAA Systems. The distinction between Provider and Deployer is a necessary first step in the deployment of either certification scheme and can be found in Section 1.0 Scope. The terms Provider and Deployer are defined specifically by both the law and in the Terms and Definitions Section 3.0.

1.0 Scope

ForHumanity designed this certification scheme for Providers (Auditees) of any size. The scheme may be applied to one or more specific AI, Algorithmic, or Autonomous Systems (including General-Purpose AI) that have been placed on the market or put into service, however it may not be used for AAA Systems that are prohibited under Article 5 of the Act. The certification scheme will cover all obligations under the EU AI Act and is valid for 12 months unless significant changes occur (see section 1.0.1 for the Audit Period of Validity).

If the AAA System is a necessary safety component of any harmonised legislation in Annex I Section A or B, and is placed on the market or put into service independently from the Product, as identified in the harmonised legislation (Recital 87), then the safety component is in scope for this certification scheme.

If the AAA System is described in Annex III, then it is in scope for this certification scheme.

If the AAA System is placed on the market or putt into service general-purpose AI models in the Union, irrespective of whether those providers are established or located within the Union or in a third country, then it is in scope for this certification scheme and will be subject to additional audit criteria based upon Chapter V of the EU AI Act.

Any AAA System, even one that is not “high-risk”, may voluntarily choose to abide by this certification scheme.

1.1.1 Determination of Provider Status

Prior to engaging with this certification scheme, it is necessary for the organisation that will be the auditee to verify that they are a Provider of the AAA System in order to use this scheme effectively.

Excerpt-only

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

A Provider of an AAA System is defined as follows:

natural or legal person, public authority, agency or other body that develops an AI system or a general-purpose AI model or that has an AI system or a general-purpose AI model developed and places it on the market or puts the AI system into service under its own name or trademark, whether for payment or free of charge;

Any distributor, importer, deployer or other third-party shall be considered to be a Provider of a high-risk AI system for the purposes of this certification scheme if any of the following apply:

1. The organisation puts their name or trademark on a high-risk AI system already placed on the market or put into service, regardless of contractual obligations with an importers, distributor, or provider of the AAA System

   or

2. The organisation makes a Substantial Modification to a high-risk AI system that has already been placed on the market or has already been put into service in such a way that it remains a high-risk AI system pursuant to Article 6 of the EU AI Act;

   or

3. The organisation modifies the intended purpose of an AI system, including a general-purpose AI system, which has not been classified as high-risk and has already been placed on the market or put into service in such a way that the AI system concerned becomes a high-risk AI system in accordance with Article 6.

Substantial Modification is defined as follows:

means a change to an AI system after its placing on the market or putting into service which is not foreseen or planned in the initial conformity assessment carried out by the provider and as a result of which the compliance of the AI system with the requirements set out in Chapter III, Section 2 is affected or results in a modification to the intended purpose for which the AI system has been assessed

To determine whether a Substantial Modification has occurred the organisation shall assess each of the Chapter III Section 2 obligations listed below to determine whether a change has been made in the AAA System

1. Risk Management (comprehensive reassessment)
2. Data Management and Governance (especially retrained General-Purpose AI models, or deviations on intended Pipeline data schema)

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

3. Change in Human Oversight (different than the required human oversight identified by the prior Provider)
4. Change in Quality Management System (reassessment of quality control and quality assurance)

The organisation shall assess the AAA System to determine whether that definition applies to the AAA System they would seek to certify.

1.1.2 Jurisdiction(s)

This certification scheme requires Providers (Auditees) to identify all applicable jurisdictions in which the AAA System operates in order to determine additional applicable legal obligations, beyond the EU AI Act. These additional legal obligations are called Relevant Legal Frameworks (a defined term). Relevant Legal Frameworks are assessed and documented in Section 5.0 criteria #6).

Establishing "if" the AAA System is in scope for this certification scheme is the first step. The scope assessment goes through the following steps to determine applicability:

1) Assess whether the Target of Evaluation (as defined in Section 1.3) falls under the definition of AI, Algorithmic, or Autonomous Systems (definitions found in section 3.0)
2) Assess whether the AAA System is a prohibited system as described in criteria #CCCC

An Auditor shall issue conditional certification for a Provider's AAA System upon demonstrating assurance of compliance with this certification scheme, pending concurrent (within a 12 month) assurance of any of the following Relevant Legal Frameworks, if applicable:

1. EU 2016/679 (GDPR),
2. EU 2019/881, (Cybersecurity)
3. EU 2019/882 (Accessibility Act)
4. EU 2022/2065 (Digital Services Act)

1.1.3 Relevant Legal Frameworks

1.1.3.1 In scope - GDPR Applicability

AAA Systems often include the use of Personal Data, therefore it is necessary to determine if the Target of Evaluation uses Personal Data in order to know whether GDPR is applicable.

Preview only

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

1.1.3.2 In scope - European Accessibility Act of 2019 (EU) 2019/882

ForHumanity and the European Union both uphold the rights of Disabled Persons according to the United Nations Convention on the Rights of Persons with Disabilities (UNCRPD). Recital 80 of the EU AI Act states "... It is therefore essential that Providers ensure full compliance with accessibility requirements, including Directive (EU) 2016/2102 of the European Parliament and of the Council and Directive (EU) 2019/882."

ForHumanity recognises that the very nature of the design and development of AAA systems is often exclusionary in the name of higher accuracy rates. Therefore, tangible remediations in support of Inclusion and Accessibility, especially accommodations, are critical to uphold the law. ForHumanity has established a certification scheme called Global Disability Inclusion & Accessibility which augments this certification scheme to assure high-risk AI systems are compliant with the law.

1.1.3.3 In scope - Digital Service (EU) 2022/2065

The Digital Services Act aims to create a safer online environment for consumers and companies in the European Union (EU), with a set of rules designed to:

1. protect consumers and their fundamental rights more effectively;
2. define clear responsibilities for online platforms and social media;
3. deal with illegal content and products, hate speech and disinformation;
4. achieve greater transparency with better reporting and oversight; and
5. encourage innovation, growth and competitiveness in the EU's internal market.

These goals match to ForHumanity's mission and as a result, ForHumanity has created the ForHumanity Digital Services Act Certification Scheme v1.1 for Covered Entities to assurance compliance with this regulation. Recital 118 states "This Regulation regulates AI systems and AI models by imposing certain requirements and obligations for relevant market actors that are placing them on the market, putting into service or use in the Union, thereby complementing obligations for providers of intermediary services that embed such systems or models into their services regulated by Regulation (EU) 2022/2065. To the extent that such systems or models are embedded into designated very large online platforms or very large online search engines, they are subject to the risk-management framework provided for in Regulation (EU) 2022/2065"

1.1.3.4 In scope - Cybersecurity minimum requirements

High-risk AI systems that have been certified or for which a statement of conformity has been issued under a cybersecurity scheme pursuant to Regulation (EU) 2019/881 and the references

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

of which have been published in the Official Journal of the European Union shall be presumed to comply with the cybersecurity requirements set out in Article 15 of this Regulation in so far as the cybersecurity certificate or statement of conformity or parts thereof cover those requirements.

ForHumanity provides integrated and harmonised certification scheme so that the auditee can be able to document satisfaction in regards to Article 15.

1. ForHumanity Cybersecurity Certification scheme
2. Regulation (EU) 2019/881- Article 54 certification - Regulation (EU) 2024/482

1.1.4 Conditions for using this certification scheme as an Importer

Importers meeting the definition in the EU AI Act may use this certification scheme, as a "Provider" by ensuring all of the following requirements before placing a high-risk AI system on the market.

importers shall ensure that the system is in conformity with this Regulation by verifying that:

1. The relevant conformity assessment procedure referred to in Article 43 has been carried out by the provider of the high-risk AI system;
2. The provider has drawn up the technical documentation in accordance with Article 11 and Annex IV;
3. The system bears the required CE marking and is accompanied by the EU declaration of conformity referred to in Article 47 and instructions for use;
4. The provider has appointed an authorised representative in accordance with Article 22(1).

In such an application all references to "Provider" are intended to mean "Importer"

1.2 Audit Period of Validity

A certification is good for one year provided additional applicable certifications (e.g., GDPR, Digital Services Act) are currently valid. Compliance should be renewed each year and an auditee is expected to maintain compliance with the current version of the audit. In any areas where the certification criteria have been changed, the auditee will have until the next annual audit to bring their systems into compliance.

Some examples of a significant change or **Substantial Modification** that require recertification to maintain status are:

1. Changes in Scope, Nature, Context, and Purpose
2. Model, Data, or Concept drift
3. Acquisition/Change in Control

Excerpt Only

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

4. Complaint(s) or Adverse Incident reports
5. Regulatory intervention
6. ForHumanity's Cause for Concern

1.3 Out of Scope Systems

AAA Systems prohibited by the Act may not be certified using this scheme.

AAA systems where and in so far they are placed on the market, put into service, or used with or without modification exclusively for military, defence or national security purposes, regardless of the type of entity carrying out those activities or AI systems which are not placed on the market or put into service in the Union, where the output is used in the Union exclusively for military, defence or national security purposes, regardless of the type of entity carrying out those activities.

AAA Systems placed on the market by Deployers are out of Scope for this certification scheme, however ForHumanity offers a separate certification scheme for them.

AAA systems operated by public authorities in a third country or international organisations, where those authorities or organisations use AAA Systems in the framework of international cooperation or agreements for law enforcement and judicial cooperation with the Union or with one or more Member States are out of scope for this certification scheme.

AAA Systems operated for the sole purpose of scientific research and development are out of scope when all of the following conditions apply:

1. There is documentable evidence of the application of the scientific method
2. There are no direct impacts to AI Subjects
3. There is no commercialisation of the AAA System, including through the monetisation of Personal Data
4. Where the live environment is not freely available

1.3 Target of Evaluation Determination Process

The organisation seeking certification determines the AAA System to which the certifying body will apply the scheme and documents this agreement in a contract. The Target of Evaluation (ToE) shall be defined by contract between the certifying body and the organisation. The certification is valid for 12 months from the date certification is issued by the certifying body.

This document is the property of ForHumanity Inc. ©2022-24, ForHumanity Inc. a 501(c)(3) tax-exempt Public Charity

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

The contract shall document all information required by the certifying body for a sufficient Certification Plan and shall include all of the following:

1) Name/identifier of the ToE, specifically noting all inputs and outputs of the **AAA System** as described in the **System Architecture Report** - Document that describes the overall, top-level blueprint of conceptual/logical/physical structure of the system including relevant frameworks and applicable standards (e.g., ISO, CEN/CENELEC, IEEE) and includes descriptions of **Processor**, and sub-**Processor** relationships including databases, processing, flow and movements, pipeline, data collection, UX interfaces, and location/**Jurisdiction** and the **Data Flow Diagram**
2) Systems or organisations expected to be "in" or "out" of scope (including a visual representation as appropriate). "In" and "out" of scope applies to third parties (including Processors) under contract.
3) The AAA System will be specifically identified including its Scope, Nature, Context, Purpose. For "out" of scope adjacent or interdependent processing or systems shown in the **System Architecture Report**, the organisation shall document and justify "out" of scope boundaries for those adjacent or interdependent processing or systems
4) Description of the data deployed in the system, specifically noting the Personal Data and Special Category Data that may be present (including Inferences and/or potential Proxy Variables)
5) Specify where the processing of data happens in terms of physical location, including whether or not there are transfers to third countries or international organisations and whether such transfers are a part of the ToE or out of scope to the ToE.
6) Identify all applicable jurisdictions in which the AAA System processes data in order to determine additional applicable legal obligations, beyond GDPR, called Relevant Legal Frameworks (as documented in Section 5.0 criteria #6).

The certifying body will only perform an audit of the documented scope. The Provider bears the responsibility of ensuring that all necessary components of the AAA System are covered in the definition of the ToE.

The ToE shall be defined in such a way that it is not misleading or likely to be misinterpreted by third parties.

The ToE may include elements of the application that are NOT AAA Systems themselves but are necessary to ensure that the AAA System functions according to the defined Scope, Nature, Context, and Purpose. This certification scheme is NOT limited to certifying only the AI, Algorithmic or Autonomous component, but rather the entirety of the AAA System application.

Excerpt Only

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

1.4 Territorial Scope

This certification scheme applies to Providers placing on the market, putting into service, or producing outputs or to the extent the output produced by AAA Systems is intended to be used in the Union, regardless of the jurisdiction of the Provider. It further applies to an AAA System that impacts EU Citizens regardless of their physical location.

5.0 Criteria catalogue

Column 1 = Reference to EU Artificial Intelligence Act articles or general oversight and accountability
Column 2 = Category
Column 3 = Audit criteria
Column 4 = Evaluation method

		Article 9 - Risk Management	
	Risk Management	In consideration of maximizing risk controls, treatments, and mitigation on behalf of both direct and indirect stakeholders including, **Protected Categories, Intersectionalities,** and **Vulnerable Populations** and in collaboration with all applicable specialty committees, the **Algorithmic Risk Committee** shall: A. Assess all **Algorithmic Risks** identified in the studies, analyses, or assessments (e.g. **Necessity, Proportionality, Algorithmic Risk, Ethical Risk**) for the **AAA System** B. Identify risk controls, treatments, and mitigations for each **Algorithmic Risk** C. Implement risk controls, treatments, and	Internal procedure manual

Exempt-only

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

		mitigations with **Traceability** D. Be accountable for the testing and evaluation process, i. Approve the **Test Plan** and **Test Completion Plan** ii. Provide specialist testing resources including an expert in testing and evaluation, designated as the **Test Lead** iii. Include the **Test Lead** as a member of the **Algorithmic Risk Committee** E. Ensure that the committee members remain trained on current testing and evaluation methods and associated outputs and reports in order to properly assess the **Test Plan** and **Test Completion Report** F. Ensure testing and evaluation includes foreseeable scenarios, capabilities, and expertise aligned and documented to appropriate to the **Test Plan** G. Ensure proportionate, based upon the risk of the **AAA System**, continuous and post-market monitoring including regularly evaluating the risk controls, treatments, and mitigations for effectiveness H. Establish and monitor an **Adverse Incidents Reporting System** I. Establish a remediation process or procedure, with **Traceability,** for when: i. Continuous or post-market monitoring identifies deviations from acceptable monitoring thresholds ii. **Adverse Incidents Reporting Systems** identifies a new or	

Excerpt only

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

		changed risk input or indicator iii. Risk controls, treatments, or mitigations that deviate from acceptable monitoring thresholds J. **Publicly** document **Residual Risk** according to **Relevant Legal Frameworks,** regulatory guidance, the **Ethical Risk Assessment**, and industry standards K. Ensure the production of the **cAIRE Report** and delivery to Top Management and Oversight Bodies	
	Risk Management	In consultation with: 1. The Enterprise/Organizational Risk Management (Chief Risk Officer) 2. The **Ethics Committee** 3. An applicable Specialty Committees and in consideration of: 1. the **Code of Ethics** 2. and **Code of Data Ethics**, the **Algorithmic Risk Committee** shall establish and document a **Risk Management Policy** for the **AAA System** that includes all of the following: A. Establishing the **AAA System** risk management framework, processes, and procedures including: i. Establishing a culture and process of constant risk assessment by all persons associated with the **AAA System** ii. Establishing a process to identify direct and indirect stakeholders (to satisfy the requirement for **Diverse Input and Multi Stakeholder Feedback**)	Internal Procedure Manual

Exemplary

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

involved in the **AAA System** who are responsible for risk identification

iii. Establishing a risk taxonomy and risk categories

iv. Establishing guidance for identifying risk inputs and indicators negatively impacting the health, safety, and fundamental rights associated with **AAA Systems**, including:
   a. **Ground Truth Availability**
   b. **Functional Correctness,**
   c. **Human Interactions**,
   d. Fairness and nondiscrimination,
   e. **Robustness**,
   f. Systemic Riskiness (e.g., **Systemic Societal Impact**)
   g. Effectiveness of embedding Ethics, Governance or Accountability structures to oversee risks.

v. Establishing a process for the evaluation of risk inputs and indicators including metrics and measurements for severity and likelihood

vi. Establishing a risk evaluation process that does all of the following:
   a. Examines existing incident reporting systems and industry-standard mitigations
   b. Examines risk indicators to

Excerpt only

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

identify root cause and negative impacts to health, safety, and fundamental rights

c. Examines risk inputs to consider a range of potential controls, treatments, and mitigations
d. Examines and tests the effectiveness of potential risk controls, treatments, and mitigations
e. Identifies metrics, measurements, and thresholds for monitoring and **Adverse Incident Reporting System** to define an **Emergent Risk** and establish processes and procedures to begin immediate risk assessment

vii. Establishing a process for assessing and implementing risk controls, treatments, and mitigations.
viii. Compiling **Residual Risks** from all risk assessments associated with the **AAA System** and document them in the **cAIRE Report**

B. Establishing metrics, measurements, and thresholds to identify excessive **Residual Risk**
C. Establishing the risk log and ensure the **Scope, Nature, Context,** and **Purpose** are defined in the **AAA Systems List**
D. Establishing the frequency for risk

Excerpt only

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

assessment, including by **Diverse Input** and **Multi Stakeholder Feedback** pool of human risk assessors in the context of the lifecycle of the **AAA System**

E. Establishing metrics, measurements, and thresholds for the following:
   i. **Severity**
   ii. **Likelihood**
   iii. **Human Interactions**
   iv. To advise Top Management and Oversight Bodies regarding the assessment of **Residual Risk** in the context of **Risk Appetite** and **Risk Tolerance**

F. Establishing metrics, measurements, and thresholds in regards to the frequency of risk reassessment:
   i. In response to a change in **Risk Input, Indicator, Severity**, or **Likelihood**
   ii. In response to inadequate risk categories, taxonomies, scales (Severity and Likelihood)
   iii. In response to a change in **Residual Risk** (comprehensive)

G. Establishing a process or procedure to reassess (Identify, Analyze, Evaluate, Treat) any of the (E.i - iii) in response to an exceeded threshold and compile a new **Residual Risk**

H. Establishing learning objectives in regards to risk management for all employees, contractors, or gig-workers, including human risk assessors, on the **AAA System** including roles, responsibilities, and duties according to the risk management policy in the context of the **AAA System**

Excerpt only

This document is the property of ForHumanity Inc. ©2022-24, ForHumanity Inc. a 501(c)(3) tax-exempt Public Charity

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

		I. Establishing the frequency of review of the **Risk Management Policy** J. Establishing the process for amending risk inputs and indicators based upon feedback from continuous and post-market monitoring, including **Adverse Incident Reporting Systems** K. Regarding all existing and newly identified risk controls, treatments, and mitigations, implementing: i. **Traceability** ii. Metrics, measurements, thresholds, and frequency of testing to determine the effectiveness L. Establishing the frequency of update to Enterprise/operational risk management with the **cAIRE Report** for the **AAA system** M. Establishing a process to receive input from the Ethical Risk Assessment to determine whether any **Residual Risk** is to be disclosed ***Note 3 - https://forhumanity.center/bok/risk-management/ provides guidance and templates for all elements of the ForHumanity Risk Management Framework***	
	Risk Management	The **Algorithmic Risk Committee** shall conduct and keep current an **Algorithmic Risk Assessment** that: A. Includes **Diverse Inputs and Multi Stakeholder Feedback** from human risk assessors B. Establishes a risk log (as described in criteria JJJJ)	Internal log, register, or database

Exemplary

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

		C. Conducts regular reviews across the lifecycle of the **AAA System** (e.g., design, development, deployment, monitoring, and decommissioning) D. Identifies risk inputs and risk indicators as negative impacts to health, safety, and fundamental rights, including: i. Reasonable foreseeable misuses ii. Insufficiencies associated with: a. **Accessibility** b. **Usability** E. Conducts **Failure Mode and Effect Analysis (FMEA)** F. Analyses risk, including identifying severity and likelihood G. Evaluates risk to identify: i. Risk controls, treatments, and mitigations, with **Traceability** of deployment ii. If applicable and appropriate, minor incident response processes and/or procedures, iii. If applicable and appropriate, major incident plans H. Compiles **Residual Risk** I. Classifies **Residual Risk** as public or confidential J. If the **AAA System** has **Personal Data**, then compiles a separate **Data Protection Impact Assessment** K. Implements specific metrics, measurements, and thresholds for risk reassessment L. Implements procedures for identifying deviations from acceptable monitoring thresholds, **Key Performance Indicators,** and **Key Risk Indicators** *[#1300]*	

Excerpt only

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

JJJJ	Risk Management	The **Algorithmic Risk Committee** shall maintain and keep current a risk log (as established in the **Algorithmic Risk Assessment**) for all identified risk inputs, indicators, controls, treatments, and mitigations identified in the **Algorithmic Risk Assessment** *[#3109]*	Internal log, register or database
	Risk Management	The **Algorithmic Risk Committee** shall ensure that, in a timely manner throughout the entire lifecycle of the **AAA System**, all risk inputs and indicators are collected and logged, in the risk log, from the following workflows: A. Risk management, B. Data management and governance, C. Testing and evaluation, D. Technical documentation E. Monitoring, F. Human Oversight G. Quality management system, including quality objective and controls H. **Adverse Incident Reporting Systems** I. Real world testing (Article 60), if applicable	Internal log, register, or database
	Risk Management	In consideration of **Residual Risk**, the **Algorithmic Risk Committee** shall implement processes and procedures to determine whether the **Residual Risk** is excessive and recommend decommissioning and document the conclusion in the **cAIRE Report**	Internal Procedure Manual
	Risk Management	In consultation with: 1. The **Ethics Committee** 2. All applicable specialty committees, Prior to deploying the **AAA System**, the **Algorithmic Risk Committee**, shall document in the **Business Rationale Report**: 3. Acceptance of the **Causal Hypothesis**	Correspondence (Internal or External)

Exceptionly

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

		4. Acceptance of the **Residual Risk** to fundamental human rights, as identified in the fundamental rights impact assessment 5. Endorsement by the **Ethics Committee** to **Place on the Market** the **AAA System** 6. Endorsement from the expert legal team (internal or external) to **Place on the Market** the **AAA System**	
	Risk Management	Using the **Scope, Nature, Context, and Purpose** of the **AAA System**, the **Algorithmic Risk Committee** shall document with **Traceability**, in the **Risk log**, a list of applicable requirements identified from the relevant EU Harmonised Standards and Common Specifications found in Annex I Section A or B	Internal log, register or database
	Risk Management	In consideration of relevant EU harmonised standards and common specification aligned to the **Scope, Nature, Context and Purpose** of the **AAA System**, the **Algorithmic Risk Committee** documents, in the **Risk Log**, risk inputs and indicators associated with conformance to the standards and specifications across the entire lifecycle of the **AAA System**	Internal log, register or database
9.2.b	Risk Management	The **Algorithmic Risk Committee** shall implement all risk controls, treatments, and mitigations with **Traceability** as documented in the **Algorithmic Risk Assessment**	Internal Procedure Manual
	Risk Management	In consideration of the Fundamental Rights Impact Assessment, the **Algorithmic Risk Committee** shall implement all identified risk controls, treatments, and mitigations with **Traceability**	Correspondence (Internal or External)

Exemplary

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

	Risk Management	The **Algorithmic Risk Committee** shall implement all risk controls, treatments, and mitigations identified in the **Proportionality Study** with **Traceability** *[#2774]*	Correspondence (Internal or External)
	Risk Management	The **Algorithmic Risk Committee** shall implement all risk controls, treatments, mitigations and/or **Ethical Choice** decisions documented in the **Ethical Risk Assessment** with **Traceability** *[#3130]*	Internal procedure manual
	Risk Management	In consideration of: 1. The **Algorithmic Risk Assessment** 2. The **Fundamental Rights Impact Assessment** 3. **Relevant Legal Frameworks**, if either assessment identified that the **AAA System** may have a discriminatory impact on any individual, group or population, then the **Algorithmic Risk Committee** shall implement the following mitigations: A. **Representativeness** testing B. Diversity and balance (stratification) testing C. Bias Mitigation And document the conclusions in the **Algorithmic Risk Assessment**	Internal Procedure Manual
9.2.b	Risk Management	If **Diverse Inputs and Multi Stakeholder Feedback** human risk assessors identify potential misuses, then the **Algorithmic Risk Committee** shall ensure that they are logged as risk inputs in the **Algorithmic Risk Assessment**	Internal Procedure Manual

Excerpt-only

This document is the property of ForHumanity Inc. ©2022-24, ForHumanity Inc. a 501(c)(3) tax-exempt Public Charity

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

9.2.b	Risk Management	In consideration of the **Relevant Legal Frameworks**, the **Ethics Committee** shall assess instances of **Ethical Choice** in the **AAA System**: A. To determine which misuses of the **AAA System** are reasonably foreseeable B. To document reasonably foreseeable misuses in an **Ethical Risk Assessment** C. To communicate the reasonably foreseeable misuses to the **Algorithmic Risk Committee**	Internal Procedure Manual
9.2.c	Risk Management	In the context of all continuous and post-market monitoring and the **Adverse Incident Reporting System**, the **Algorithmic Risk Committee** shall ensure that all incidents are identified as risks (either as a risk input or indicator) and: A. Classified as security or algorithmic risk B. Logged as risk inputs or indicators in the Cybersecurity Risk Log or the Algorithmic Risk Log as applicable C. Analysed, evaluated and treated according to the applicable policy	Internal Log, Register, or Database
9.2.c	Risk Management	In consideration of: 1. Continuous and post-market monitoring 2. **Adverse Incident Reporting System**, 3. **Key Risk Indicators (KRIs)** the **Algorithmic Risk Committee** shall assess incidents (either as a risk input or indicator) to determine whether they exceed established thresholds indicating a full risk reassessment, and document the conclusion in the Algorithmic Risk Assessment, including lessons learned from exceeded KRIs.	Internal procedure Manual
	Risk Management	In consideration of: 1. Continuous and post-market monitoring 2. **Adverse Incident Reporting System**,	Internal procedure Manual

Excerpt only

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

		the **Algorithmic Risk Committee** shall assess to determine whether any newly identified risks (either as a risk input or indicator), exceed the thresholds indicating comprehensive risk reassessment and document the conclusions in the **Algorithmic Risk Assessment**	
9.2.c	Risk Management	If the root cause of a risk indicator can be identified through analysis and evaluation of the risk by the **Algorithmic Risk Committee** then the risk indicator shall be changed to a risk input and treated accordingly else, the failure to determine a root cause shall be included in the **Residual Risk** and the risk indicator shall be treated accordingly	Internal Procedure Manual
9.2.d	Risk Management	The **Algorithmic Risk Committee** shall: A. Assess all risk inputs and indicator to determine the most effective risk controls, treatments, and mitigations that collectively minimises **Residual Risk** B. Implement the risk controls, treatments, and mitigations with **Traceability** C. Implement a monitoring process with metrics, measurements, and thresholds for each risk control, treatment, and mitigations to determine continued effectiveness	Internal Procedure Manual
	Risk Management	In consideration of the: 1. **Test Completion Report**, 2. **Algorithmic Risk Assessment**, 3. **Ethical Risk Assessment**, 4. **Data Curation Report** 5. **Data Evaluation Report** 6. **Relevant Legal Frameworks** and regulatory guidance 7. **Provider's Risk Appetite and Risk Tolerance**, the **Algorithmic Risk Committee** shall accept and compile all individual unmitigated risks in	Internal Procedure Manual

Exposure Draft

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

		the following manner: A. In manner that meets all legal obligations B. A risk controlled, treated, or partially mitigated but not entirely eliminated establishes an individual unmitigated risk C. Each individual unmitigated risk must be documented and accepted by the **Algorithmic Risk Committee** or the **Ethics Committee** (for instances of **Ethical Choice)** D. All individual unmitigated risks must be combined to establish the AAA System's **Residual Risk** and document the conclusions in the **cAIRE Report**	
	Risk Management	In consideration of the established thresholds for public disclosure found in the **Ethical Risk Assessment,** the **Algorithmic Risk Committee** shall establish a procedure to assess whether the unmitigated risk exceeds the thresholds to determine if the unmitigated risk is to be disclosed **Publicly** and document the conclusion in the **cAIRE Report**	Physical Testing
	Risk Management	In consideration of: 1. **Relevant Legal Frameworks** 2. Regulatory Guidance 3. The threshold(s) for confidential disclosure to National Supervisory Authorities In regards to each unmitigated risk, the **Algorithmic Risk Committee** shall assess to determine whether the unmitigated risk exceeds the thresholds indicating that the risk is to be disclosed confidentially to National Supervisory Authorities and document the conclusion in the **cAIRE Report**	Physical Testing

Excerpt-only

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

	Risk Management	In consideration: 1. The **Residual Risk** conclusion in the correspondence from Top Management and Oversight Bodies 2. **Relevant Legal Frameworks** 3. Regulatory guidance the **Algorithmic Risk Committee** shall assess the **Residual Risks** to be disclosed **Publicly** to determine the language and medium in which they are to be compiled, documented, and displayed	Public Disclosure Document
LLLL	Risk Management	In consideration of the accepted **Residual Risk**, the **Algorithmic Risk Assessment**, the **Algorithmic Risk Committee** shall assess the severity and likelihood of each individual **Residual Risks** to determine which individual risks are to to be disclosed in the **Just-in-Time** notification or the **Terms and Conditions** and document conclusions in the **Algorithmic Risk Assessment**	Internal Procedure Manual
	Risk Management	Based upon the assessment in Criteria LLLL, the **Algorithmic Risk Committee** shall ensure that the public disclosure of **Residual Risk** is disclosed in a **Just-in-Time"** notification using clear and plain language containing all of the following: A. Specific unmitigated individual risks based upon a risk-based assessment of severity and likelihood to **Deployers/ AI Subjects** B. That the **AI Subject** is interacting with an **AAA System** C. The known level of **Functional Correctness** along with **Explainability** and if applicable, **Explainability+ Statement** D. An adequate disclaimer for **Adverse Impacts** and directions on how to access	Physical Testing

Exemplary

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

		the **Adverse Incident Reporting Systems** for **Deployers/AI Subjects** E. An opportunity for **Deployers/ AI Subjects** to opt-out as far as feasible F. A conspicuous link to the **Terms and Conditions**	
	Risk Management	Based upon the assessment in Criteria LLLL , the **Algorithmic Risk Committee** shall document public **Residual Risk** in the **Terms and Conditions**	Physical Testing
	Risk Management	If the **Ethics Committee** assesses the **AAA System** to be **Novel**, then the **Algorithmic Risk Committee** shall ensure all of the following: A. Specific disclosure in the **Residual Risk** that the **AAA System** is **Novel** documented in the **risk log** with **Traceability** B. Augment training to sales, marketing and promotional teams, with **Traceability**, on the added risks associated with a **Novel AAA System** to avoid false, misleading or exaggerated claims C. Augment disclosures in the **AAA System Deployer's or AI Subject's Guide** stating the lack of industry standards and sufficient comparables	Correspondence (Internal or External)
	Risk Management	In consideration of: 1. the **Ethical Risk Assessment** 2. The metrics, measurements, and thresholds provided by the **Ethics Committee** in regard to the health, safety, and well-being of human interactors, the **Algorithmic Risk Committee** shall establish a process or procedure to measure and monitor human interactors with **Traceability** to	Correspondence (Internal or External)

Exception only

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

		the **Ethics Committee**	
	Risk Management	In consideration of the **Scope, Nature, Context, and Purpose** and the technical infrastructure of the **AAA System**, the **Algorithmic Risk Committee** shall establish the expected lifetime of the system and document it in the **AAA Systems List,** and the **AAA System Deployer Guide**	Internal Procedure Manual
	Risk Management, Recital 118	If the **AAA System** is governed by the Digital Services Act (EU 2022/2065), then the **Algorithmic Risk Committee** shall ensure that **Diverse Input and Multi stakeholder Feedback** human risk assessors include risk inputs and indicators that are not covered by the Digital Services Act in the **Algorithmic Risk Assessment**	Internal Procedure Manual

Appendix A - Infrastructure of Trust for AI - Guide to Entity Roles and Responsibilities

ForHumanity promotes this certification scheme to all entities that wish to provide advice, guidance, consulting and assurance or organisation both outside and within the European Union. ForHumanity licences entities to offer the scheme and a list of licensed entities can be found here.

ForHumanity also trains individuals to become ForHumanity Certified Auditors (FHCAs). Earning this certification is the ultimate assurance of knowledge of this certification scheme and the process by which certification is achieved for organisations. ForHumanity offers online, asynchronous training in this certification scheme through its training platform – ForHumanity University.

ForHumanity promotes this and many other certification schemes to organisations, governments, regulators, national accreditation bodies, professionals and the public by social media, conference speeches, university lectures, online presence, and execution of our mission

Excerpt-only

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

statement to specific support an infrastructure of trust with a wide range of participants from society.

Describing the roles in an infrastructure of trust for AI, Algorithmic and Autonomous (AAA) Systems - we have a model with a long track record of success. ForHumanity is adapting that model to AAA Systems.

Background on Independent Audit

In 1973, the accounting industry came together and formed The Financial Accounting Standards Board (FASB) which created the Generally Accepted Accounting Principles (GAAP) which still govern financial accounting today. Eventually, the US Securities and Exchange Commission, and other extranational regulatory agencies, required adherence to the GAAP standard for all publicly listed companies. This clarity and uniformity significantly improved the financial world. An infrastructure of trust has been built over the past 50 years because of critical features such as independence, certified practitioners, and third-party rules that are compliant with the law and best practices.

Adapting to AI and Autonomous Systems

ForHumanity has advocated for the adoption of this infrastructure of trust and explained how it can be adapted and adopted for the Governance, Accountability, and Oversight of AI and Autonomous Systems. We support the creation and mandate of Independent Audit of AI Systems (IAAIS). IAAIS provides a comprehensive solution grounded in the same fundamental principles as Independent Financial Audit. ForHumanity develops and maintains audit and certification criteria designed for a range of industries and jurisdictions.

The proposed system replicates the distributed oversight, accountability and governance needed for AI, Algorithmic, and Autonomous (AAA) systems in the same manner as financial audit, through audit and pre-audit service providers. These entities will employ certified practitioners to prepare for an eventual independent audit performed by other certified practitioners. The audit criteria are crowdsourced and presented transparently to maximise an entity's ability to achieve compliance. Advancements in systems technology allow many of these processes to be automated for entities such as with the Treadway Commissions' Committee of Sponsoring Organization (COSO) framework for internal risk, audit and controls. The result is a fully-integrated, compliance-by-design infrastructure that embeds human agency, transparency, disclosure and compliance from design to decommission.

Role on Independent Audit of AI and Autonomous Systems

The audit criteria are applied in two vectors: 1) Top-down accountability, governance and oversight 2) laterally, AI system by AI system. The top-down approach creates accountability systems for ethics, bias, privacy, trust, and cybersecurity for the Board of Directors, Chief Executive Officer, and Chief Data Officer. Committee structures are required such as an Algorithmic Risk, Ethics and specialty committees to manage the audit/compliance responsibilities, as a second line of defence. All of these top-down criteria apply to every AI and every autonomous system in the organisation. The system-specific audit criteria are designed to ensure legal and best practice compliance tailored to the specific impact of each system on humans. This comprehensive approach ensures consistency across the organisation combined with complete risk management coverage of each unique system.

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

Participants in the System

The roles largely remain the same in Independent Audit of AI Systems as described in Taxonomy. There are six distinct roles in most jurisdictions. Each player performs their function and the rules are executed in the same conflict-free manner, ensuring the highest integrity.

Certifying Bodies/Notified Bodies/Auditors (Auditors)

- An Auditor engages in 3-party contract party contracts, with the Target of Evaluation (ToE) and on behalf of the public or intended users.
- The auditor deploys certified practitioners to conduct the audits.
- The auditor itself is certified by the Government Accreditation Service.
- When audits are conducted there is no feedback loop to the company and the audit is compliant or non-compliant.
- Audits are publicly disclosed according to the rules of the jurisdiction.
- The Auditor is liable for false assertions of compliance
- An Auditor is licensed for use of certification criteria
- The Auditor shall not provide Pre-audit services to Audit clients
- An Auditor may provide Pre-Audit services to non-Audit ToEs (may require accreditation)

Pre-Audit Service Providers/Consultants/Advisors (PASP)

- PASP engages in a 2-party contract directly with the Target of Evaluation
- There is a direct feedback loop between the ToE and PASP
- The PASP may or may not deploy certified practitioners per local jurisdiction rules
- The PASP may or may not be accredited by the Government Accreditation Service
- The PASP offers no certification or guarantee of audit compliance
- The PASP works are private, on behalf of the ToE
- The PASP is not liable for failed compliance or false assertions of compliance
- The PASP may or may not be licensed for use of certification criteria, but must be licensed if the service offered is related to or designed to satisfy certification requirements
- The PASP shall not be the auditor for a PASP client
- A PASP may offer Audit service to non-PASP clients (must be accredited)
- A PASP may deploy compliance-in-a-box solutions for criteria compliance

Entities seeking Certification/Providers/Deployers (Auditee)

- Auditee may engage PASP
- Auditee shall have an Auditor if required by the Relevant Legal Framework
- Auditee pledges that all components, systems and relevant, supporting infrastructure to be certified will be disclosed to the Auditor, failure in this regard is the responsibility of the ToE
- Auditee dealings with PASP shall be confidential and non-public audit compliance may be confidential with an Auditor
- Auditee shall maintain compliance structures, such as Algorithmic Risk Committee, Children's Data Oversight Committee, and Ethics Committee

This document is the property of ForHumanity Inc. ©2022-24, ForHumanity Inc. a 501(c)(3) tax-exempt Public Charity

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

- Auditee shall build and maintain internal controls and systems to aid in compliance with audit requirements and foster robust risk management, monitoring, and regulatory compliance
- Auditee shall be responsible for all public disclosures

Third-Party Criteria creation, maintenance, and individual certifier (ForHumanity)

- Non-profit organisation
- Independent of Auditors and PASP
- Transparent and inclusive of input and critique from all participants
- Criteria designed to uphold human well-being
- Conflict-free of undue Auditee influence
- Submits to the authority of the jurisdiction for certified criteria
- Iterates and maintains criteria consistent with the law and best practices in a binary and auditable fashion
- Trains and certifies individual practitioners on all criteria in support of uniformity of audit assurance process
- Maintains a transparent repository of use cases and knowledge stores in support of Auditors/Auditees to facilitate compliance
- Licences criteria to all qualified Certifying Bodies/Notified Bodies/Auditors/PASP
- Provides standard contract clauses for Auditors and PASP
- Engages in distributed education system to maximise availability and certified individuals
- Maintains a system of Continuing Education (CE)
- Maintains a searchable, registration system of Accredited Individuals and holds them to a Code of Ethics and Professional Conduct
- Ensures Independence and anti-collusion amongst of Certifying Bodies/Notified Bodies/Auditors/PASP
- Maximises global harmony amongst audit criteria while ensuring jurisdictional sensitivity

Government-approved Accreditation Service

- Creates trust and confidence in products and services
- Assures that Certified/Notified/Accredited Bodies have sufficient talent, skill, scope, and financial foundation to provide certification
- Regular review of accreditation standards
- Regular review of Certified/Notified/Accredited Bodies
- Regular review of Third-party Criteria provider and individual certification
- Determines form and elements of Post Audit Compliance Report
- Maintains an accessible list of Certified/Notified/Accredited Bodies
- Maintains an accessible list of sanctioned or suspended Certified/Notified/Accredited Bodies

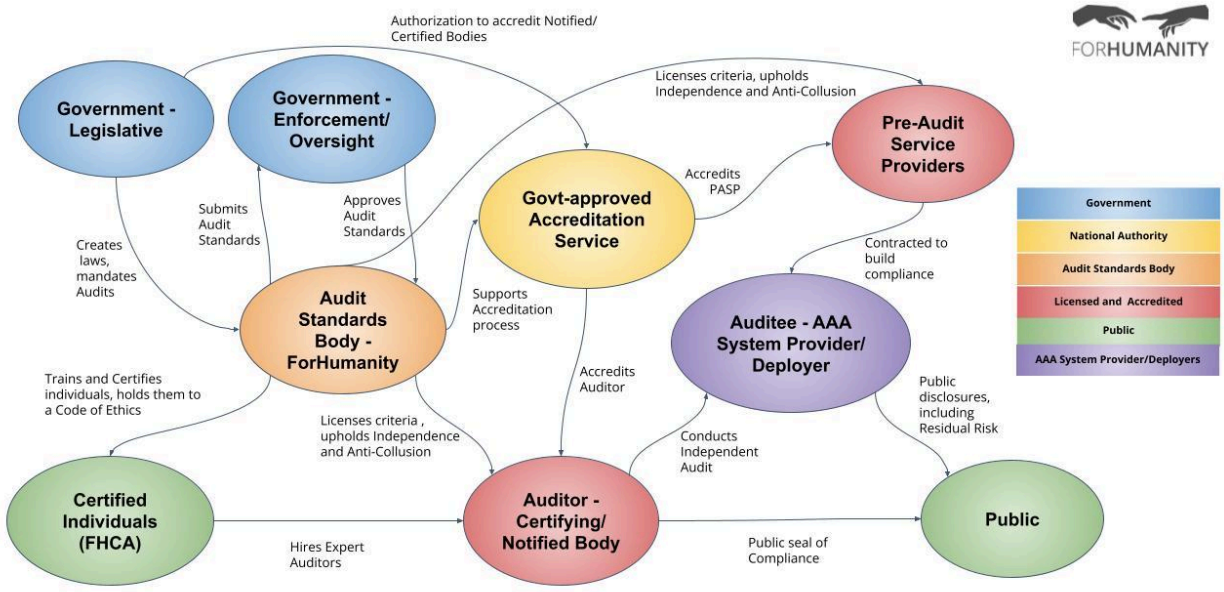
Governments/Regulators or similar Law-making/enforcement body

- Democratically, elected body
- Legislative responsibilities
- Executive or enforcement responsibilities
- Establishes prohibited AAA Systems

This document is the property of ForHumanity Inc. ©2022-24, ForHumanity Inc. a 501(c)(3) tax-exempt Public Charity

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

- Establishes low risk and exclusionary criteria from mandatory Independent Audit
- Regularly meets to review laws and best-practices
- Establishes a panel of experts to reviews and accredits (or rejects) submitted criteria
- Engages in enforcement actions for non-compliance with the law
- Handles concerns and issues brought by the Public



Roles and Responsibilities - Independent Audit - Infrastructure of Trust

Licensing

ForHumanity provides four types of licences:

A. Auditor/Certification Body and Pre-Audit Service Provider
B. Platform, technology, or SaaS tools
C. Teaching (for commercial purposes)
D. University (for academic and research purposes) as well as commercial use of certification course

Any entity that uses the certification scheme as the basis of their business relationship (generating revenue or a similar quid pro quo - commercial purposes) with a client must be duly licensed. Any organisation may be licensed by ForHumanity, but they must also have FHCAs on staff in good standing to issue certificates or provide services using the intellectual property.

Audit fees are owed upon receipt of revenue by a licensee. The licence fees allow ForHumanity to maintain the certification schemes and training individuals as experts or ForHumanity

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs

Certified Auditors (FHCA). Trademarks, certification marks, audit criteria, and services marks of ForHumanity are provided in licensing agreements and must be used in adherence with the terms of service found in the licence agreement. All licence agreements contain identical terms and conditions as relatable across use cases and are non-negotiable to ensure uniformity.

Excerpt-only

All rights reserved. Creative Commons CC-BY-NC-ND Attribution-NonCommerical-NoDerivs