

IEEE CertifAIEd™ – Ontological Specification for Ethical Privacy

Abstract: The IEEE CertifAIEd™ criteria for certification in ethical privacy are discussed in this ontological specification. Providing actionable methods to granularly assess and benchmark systems and organizations in their ethical performance is the goal of this work. Original methods of analyzing the respective drivers and inhibitors that influence the emergence of a quality of ethics, in this case privacy, are utilized by the certification methodology. The creation of the certification process is discussed, along with its intended implementation. An overview of the criteria schema and example criteria are also provided. This certification process has been designed to generate a tailorable and scalable system for the development of conformity assessment and certification for emergent ethical features of autonomous intelligent systems (AIS). The contents of this ontological specification are designed to be broadly applicable to a wide variety of domains and use-cases as well as providing flexibility through up to three levels of criteria, enabling a deeper and more sophisticated scrutiny and certification process where necessary.

Keywords: autonomous intelligent systems, ethics, privacy

The Institute of Electrical and Electronics Engineers, Inc.
3 Park Avenue, New York, NY 10016-5997, USA

IEEE is a registered trademark in the U.S. Patent & Trademark Office, owned by The Institute of Electrical and Electronics Engineers, Incorporated.

IEEE CertifAIEd™ is a trademark owned by The Institute of Electrical and Electronics Engineers, Incorporated.

IEEE prohibits discrimination, harassment, and bullying.
For more information, visit <https://www.ieee.org/about/corporate/governance/p9-26.html>.



This Work is licensed under an [Attribution-NonCommercial-NoDerivatives 4.0 International License \(CC BY-NC-ND 4.0\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

TRADEMARKS AND DISCLAIMERS

IEEE believes the information in this publication is accurate as of its publication date; such information is subject to change without notice. IEEE is not responsible for any inadvertent errors.

The ideas and proposals in this specification are the respective author's views and do not represent the views of the affiliated organization.

Notice and Disclaimer of Liability Concerning the Use of IEEE SA Documents

This IEEE Standards Association (“IEEE SA”) publication (“Work”) is not a consensus standard document. Specifically, this document is NOT AN IEEE STANDARD. Information contained in this Work has been created by, or obtained from, sources believed to be reliable, and reviewed by members of the activity that produced this Work. IEEE and the IEEE Conformity Assessment Program (ICAP) members expressly disclaim all warranties (express, implied, and statutory) related to this Work, including, but not limited to, the warranties of: merchantability; fitness for a particular purpose; non-infringement; quality, accuracy, effectiveness, currency, or completeness of the Work or content within the Work. In addition, IEEE and the ICAP members disclaim any and all conditions relating to: results; and workmanlike effort. This document is supplied “AS IS” and “WITH ALL FAULTS.”

Although the ICAP members who have created this Work believe that the information and guidance given in this Work serve as an enhancement to users, all persons must rely upon their own skill and judgment when making use of it. IN NO EVENT SHALL IEEE SA OR ICAP MEMBERS BE LIABLE FOR ANY ERRORS OR OMISSIONS OR DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO: PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS WORK, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE AND REGARDLESS OF WHETHER SUCH DAMAGE WAS FORESEEABLE.

Further, information contained in this Work may be protected by intellectual property rights held by third parties or organizations, and the use of this information may require the user to negotiate with any such rights holders in order to legally acquire the rights to do so, and such rights holders may refuse to grant such rights. Attention is also called to the possibility that implementation of any or all of this Work may require use of subject matter covered by patent rights. By publication of this Work, no position is taken by the IEEE with respect to the existence or validity of any patent rights in connection therewith. The IEEE is not responsible for identifying patent rights for which a license may be required, or for conducting inquiries into the legal validity or scope of patents claims. Users are expressly advised that determination of the validity of any patent rights, and the risk of infringement of such rights, is entirely their own responsibility. No commitment to grant licenses under patent rights on a reasonable or non-discriminatory basis has been sought or received from any rights holder.

This Work is published with the understanding that IEEE and the ICAP members are supplying information through this Work, not attempting to render engineering or other professional services. If such services are required, the assistance of an appropriate professional should be sought. IEEE is not responsible for the statements and opinions advanced in this Work.

Participants

At the time this ontological specification was completed, the IEEE CertifAIEd™ Privacy Expert Working Group had the following membership:

Patricia Shaw, *Chair*
Ali Hessami, *Technical Editor*

Elenor (Nell) Watson
Gerlinde Weger

Ali Hessami

Patricia Shaw
Scott L. David

The Privacy Expert Focus Group

The work of IEEE CertifAIEd™¹ was largely driven by the efforts of expert focus groups, their appointed leads, and support from the Chair. The Privacy Expert Focus Group (PEFG) was formed of volunteers from diverse backgrounds and experience, including legal, computer science, technological, organizational, safety, auditing, and fiscal. However, other experts were invited to complement gaps identified in the profile of PEF. The PEF held 16 ideas capture workshops in developing the ethical privacy schema, a graphical representation of factors that positively or negatively influence ethical accountability, which is set out in Annex A.

¹ IEEE CertifAIEd™ is a trademark owned by The Institute of Electrical and Electronics Engineers, Incorporated.

Introduction

The advent of automation during the industrial revolution brought about societal and business benefits in large-scale production, consistency, quality, and efficiencies that made commodities affordable. One key feature of most automation systems is the existence of human in the loop (HITL) at some stage providing oversight and control on critical aspects of the process or production. The development of *learning* machines that can perform specific tasks without using explicit instructions is now the foundation of autonomous intelligent systems (AIS) proliferating pervasively in all facets of industry, service provision, and governance. These machines rely on patterns and inductive or deductive inference, thereby raising the prospect of autonomous decision-making (ADM) by algorithmic learning systems (ALS), or ADM/ALS.

ADM/ALS offers the possibility of reducing and ultimately removing the human agent from operation, control, and supervisory roles, thereby reducing costs and potential errors while processing a much larger number of transactions offering higher service levels. While this brings savings, efficiencies, and business benefits, the removal of the human agent from the control and oversight loop brings about uncertainties and concerns regarding trustworthiness, fairness, explicability, and rationality of the automated decisions.

The uncertainties and societal concerns over ethicality and trustworthiness of ADM/ALS in all walks of life, especially in high-risk environments such as transportation, healthcare, financial, and public services, pose a formidable challenge to the uptake and innovation in deployment of the AIS-based solutions. There is thus a desire to regulate the implementation of ADM/ALS in order to provide a safety net and assurance about potential risks and societal harms that may ensue in the course of pursuing the perceived benefits.

From a broader ethical perspective, key areas of concern in development and deployment of ADM/ALS relate to accountability, transparency, freedom from unacceptable algorithmic bias/fairness, privacy, and responsible governance. To this end, the IEEE Standards Association (SA) has developed a suite of detailed criteria for evaluation, conformity assessment, and certification of these properties of ADM/ALS products and services through CertifAIEd™. This program is a key facet of the IEEE SA's Global Initiative and Ethically Aligned Design portfolio.

Contents

1. Overview	6
1.1 Scope	6
1.2 Purpose	6
2. Definitions, acronyms, and abbreviations	6
2.1 Definitions	6
2.2 Acronyms and abbreviations	7
3. Stakeholders	7
4. Context	8
5. Ethical privacy factors	8
5.1 Drivers of ethical privacy	8
5.2 Inhibitors of ethical privacy	9
6. Ethical privacy certification criteria	10
6.1 Privacy ethical foundational requirements (EFRs)	10
6.2 Normative and instructive privacy EFRs	10
6.3 Duty holders of the privacy EFRs	11
6.4 The levels of ethical privacy certification	11
6.5 Required evidence	12
6.6 Evaluation of evidence	12
6.7 The constraints of ethical privacy certification	12
Annex A AIS ethical privacy schema	13
Annex B Ethical privacy certification criteria	14
Annex C Bibliography	25

1. Overview

1.1 Scope

The IEEE ethics certification criteria developed for assurance of many ethical facets of the development and deployment of autonomous intelligent systems (AIS) constitute an extensive hierarchical suite, developed by a panel of competent experts through a model-based creative process. The criteria suite for ethical privacy comprises articulation of pertinent critical factors at two levels of hierarchy: Level 1 and Level 2. The Level 1 and Level 2 criteria collectively constitute the entire ethical privacy suite for the purposes of conformity assessment and certification. This ontological specification provides insight into and specification of Level 1 ethical privacy factors to disseminate and enhance the understanding of IEEE's ethics certification criteria.

The ethics criteria are also developed from a general ethics perspective. The development strategy and deployment approach for these criteria provide an efficient and pragmatic approach for customization of a given suite for application-specific context and requirements. This is referred to as *profiling* and, in practice, the generic ethical privacy suite can be customized into many profiles appropriate to the requirements, terminology, context, and priorities of a given sector, culture, or application vertical. This specification examines the generic ethics for ethical privacy.

1.2 Purpose

This ontological specification discusses the development and specification of ethical privacy conformity assessment and certification criteria of IEEE CertifAIED™¹. The criteria are applicable to all ethical privacy concerns within the context of AIS.

2. Definitions, acronyms, and abbreviations

2.1 Definitions

For the purposes of this document, the following terms and definitions apply.

ethical privacy: A contextual set of values pertaining to privacy and the satisfaction of a framework of expectations (preservation of autonomy, self-determination, and self-selected communities/locum and intimacies).

NOTE 1—Various dimensions—such as geographic, cultural, and ethnic—are relevant.

NOTE 2—Principles, ethics, and norms inform the paradigm.

NOTE 3—Ethics is human focused, so ethical privacy is human centric/anthropomorphic.

NOTE 4—Norms describe right and wrong actions that lead to judgments of good or evil persons or actions made by or on behalf of persons.

¹ IEEE CertifAIED™ is a trademark owned by The Institute of Electrical and Electronics Engineers, Incorporated.

NOTE 5—Ethical privacy overlaps with, and is largely complementary to, the aspects enforced and protected by law.

NOTE 6—It is recognized that in some common-law jurisdictions, the common law of privacy comprises the following torts: peeping Tom, publication of private facts, defamation, and misappropriation.

NOTE 7—The inner sphere pertains to ethical privacy being in the physical, the online, and one’s thought life or, more simply, informational privacy and data protection concerns.

NOTE 8—While the legal constructs of privacy exist, this schema is intended to reach into wider aspects of privacy, including privacy being the inner sphere of life and the public identity of an entity (individual, group, community) upholding dignity.

NOTE 9—There is a distinction (and some overlap) between privacy and autonomy.

NOTE 10— Expectations and integrity of self are the focus.

NOTE 11— Issues of human dignity and dependency in the use of technology are pertinent.

2.2 Acronyms and abbreviations

ADM	autonomous decision-making
AIS	autonomous intelligent system(s)
ALS	algorithmic learning system
EFR	ethical foundational requirement
ML	machine learning
PII	personal identifiable information

3. Stakeholders

The key stakeholders of the ethical privacy of autonomous intelligent systems (AIS) are the following entities: developers, system/service integrators, system/service operators, maintainers, regulators, and the end users (see 6.3 on duty holders).

NOTE 1—An entity can be an individual, a single organization, or a group of collaborating individuals and organizations. The above labels for the five groups of stakeholders are generic and can be mapped in terms of activities and influence against the life cycle but with overlapping activities. A single entity may assume multiple roles, that is, a developer may also fulfill and complete system design, integration, and maintenance.

NOTE 2—End users are a legitimate class of stakeholders, but there are no requirements placed on this group in these criteria.

4. Context

The IEEE CertifAIEd™ has been designed to generate a tailorable and scalable system for the development of conformity assessment and certification for emergent ethical features of AIS. This program developed ethical criteria for transparency, accountability, and algorithmic bias during an earlier phase. The current focus is on ethical privacy criteria that go beyond legal stated requirements of privacy and complement the legally enforceable protection measures. During explorations, it became clear how multifaceted and complex the issue of privacy is and how it extends beyond the notion of compliance with privacy as currently denoted in the law. Also noteworthy is that not all jurisdictions approach privacy in their respective legal systems in the same way; therefore, there was more of a need to identify this suite of criteria to help organizations assess and conform to ethical privacy.

At the commencement of the exploratory and creative approach to the development of the principal concepts and formulation of the criteria, privacy and ethical privacy were broadly defined as in 2.1.

As AIS are increasingly interwoven in human daily existence, the risk of intrusion increases in often unknown and insidious ways. The private spheres of life and the public identity of individuals, groups, and communities may be compromised, along with their denizens' dignity and opportunity for human flourishing. With respect to AIS, special attention is warranted because AIS have an ability—superior to that of any human or human organization—to glean insight from vast amounts of data. As a result, AIS have the potential to warp human input and output channels in ways that humans (individually and in groups) may not be able to defend.

As such, the IEEE CertifAIEd™ ethical privacy criteria suite comprises a holistic and systemic set of factors required in decision-making, rulemaking, enforcement, redress, operational governance, and, most importantly, human capacity and behavior across not only the AIS life cycle but with assumptions and dependencies from the wider AIS ecosystem as well. The criteria have also sought to emphasize the importance of contextual understanding, culture, and continuous monitoring to ensure appropriateness and timeliness of interventions. Furthermore, for the purposes of accountability, this suite of ethical criteria reflects an effort to have responsibility remain with the humans and human organizations involved in the actions bringing AIS into being as it is still considered premature to preassign any such responsibilities to the AIS themselves.

5. Ethical privacy factors

In considering what goals/factors contribute to the quality of ethicality—in addition to the classical identification of contributory factors—we recognized a need, supported by the adopted methodology, to map those goals/factors that would detract from it also. These are referenced as *drivers* and *inhibitors*, respectively, in the privacy schema (see Annex A). The rationale being many real-world constraints can frustrate well-meaning objectives due to issues of human resourcing, management, technological limitations, and cultural change.

5.1 Drivers of ethical privacy

The seven supportive influencing factors (drivers) impacting ethical privacy are the following:

- a) *Organizational governance, capability, and maturity*: This driver goal deals with the organization's capability, maturity, governance processes, and political will/good faith for ethical privacy assurance.

- b) *Clarity and consistency of AIS operations*: This driver goal seeks to ascertain a clear definition and the articulation and communication of the concepts and results of operation in the intended environments for AIS products, services, or systems to the relevant stakeholders.
- c) *Ethical architecture, design, and development for AIS*: This driver goal identifies whether an organization is upholding a holistic approach to ethical design and development at all levels, empowering staff to review the activities and focus of peers and provide feedback. This goal includes having due regard (being holistic, consultative, and providing for feedback) for all attributes and aspects of the architecture, design, and development that could be invasive to ethical privacy and inimical to fulfilling the ethical privacy requirements.
- d) *Human oversight and enforcement in AIS*: This driver goal identifies the human oversight involved. Human agents should be able to understand an AIS product, service, or system behavior in order to be able to intervene if necessary, to establish a process to cease activity, and to assess the context to ensure timely corrective action. In implementing human oversight, the organization should be mindful of and mitigate against harmful or detrimental types of intervention, including the risks to ethical privacy due to human oversight.
- e) *End-user awareness of AIS and empowerment*: This driver goal seeks to ascertain how potential users are being made aware of the existence and functions of an AIS element within products, services, or systems in the context of use and how they are being empowered to sufficiently understand and make decisions of the use of such systems. This may also identify where there is a disadvantage to the end user due to a lack of suitable alternative options.
- f) *Maintaining ethical privacy integrity*: This driver goal looks at efforts to maintain an ethical profile of AIS products, services, or systems with respect to privacy requirements and criteria/behaviors across the AIS life cycle and beyond.
- g) *Decommissioning*: This driver goal considers the risk and control mechanisms put in place in the decommissioning of AIS. Such processes may concern data (anonymization/deletion), metadata, insight and inference, learning and legacy code, or models. Decommissioning may also put the end user at a disadvantage due to lack of suitable alternative options.

5.2 Inhibitors of ethical privacy

The five constraining influencing factors (inhibitors) impacting ethical privacy are as follows:

- a) *Overreaching and overfitting*: This inhibitory goal relates to the use of technologies that overstep the bounds of dignity or appropriateness by either overfitting of certain characteristics or drawing unreasonable inferences based upon isolated data points. This could include unwarranted and unexpected (from the user's perspective) cross-correlation of data sets.

- b) *Authoritarian and compulsory pressures*: This inhibitory goal considers the demands and the ability by some institutions or governmental bodies to gain access to information on AIS as held by a given organization/duty holder.
- c) *Accidental/incidental exposure*: This inhibitory goal considers the inadvertent/unintentional loss or breach of security and loss of control of AIS (data, the system, and the platform), including eavesdropping and acquisition or interception of downstream data likely to compromise privacy.
- d) *Malicious exposure*: This inhibitory goal considers the intentional breach of security and loss of control of system/data or unauthorized access to the data/system, including eavesdropping and acquisition or interception of downstream data likely to compromise privacy.
- e) *Systemic vulnerability*: This inhibitory goal relates to the structural stochasticity in the AIS learning system that can pose a risk to or undermine privacy. This may appear in different components and may not be a permanent state of AIS. Any time an algorithm is transferred from one system to another, this phenomenon may be encountered.

Explanation of the goals and associated requirements, requisite evidence, and scale of measurement are depicted in Annex B.

6. Ethical privacy certification criteria

6.1 Privacy ethical foundational requirements (EFRs)

The ethical privacy schema, in conjunction with the privacy ethical foundational requirements (EFRs), enables the auditing of organizations and their autonomous intelligent technologies for ethical privacy with clear criteria that can be turned into a scoring mechanism. As a model-based approach, the schema captures both negative and positive aspects (inhibitors and drivers, respectively) of ethical privacy for AIS with ease of reference. It represents an efficient means of real-time creative knowledge capture as well as operating as the foundation for development of ethical privacy requirements.

The detailed privacy EFRs are depicted in Annex B.

6.2 Normative and instructive privacy EFRs

The privacy EFRs contain a series of expected behavioral norms and instructions on how to enact aspects of the certification, without going into specifics where not strictly necessary, in order to preserve flexibility of implementation within a bounded set of principles. In this spirit, the privacy EFRs depicted in Annex B are classed into *normative* (mandatory) and *instructive* (recommended) for the purposes of conformity assessment against the suite of ethical privacy certification criteria.

6.3 Duty holders of the privacy EFRs

The privacy EFRs depicted in Annex B are additionally noted against the specific group of duty holders for the purposes of conformity assessment. The principal groups are as follows:

- *Developer (D)*: The entity (see NOTE 1— Clause 3) that designs and develops a component (product) or system for a general or specific purpose/application. This could be as a result of a developer’s own instigation or response to the market or a client requirement. The developer is responsible for the ethical assurance of the generic or application-specific product or system and associated supply chain.
- *(System/service) Integrator (I)*: The entity that designs and assures a solution through integrating multiple components, potentially from different developers, and tests, installs, and commissions the whole system in readiness for delivery to an operator. The system delivery may take place over several stages. The integrator is usually the duty holder for total system assurance and certification, safety, security, reliability, availability, sustainability, and so forth. For this, it may rely on the certification or proof of ethics from various developers or the supply chain.
- *(System/service) Operator (O)*: The entity that has a duty, competences, and capabilities to deliver a service through operating a system delivered by an integrator.
- *Maintainer (M)*: The entity tasked with conducting required monitoring, preventive or reactive servicing and maintenance, and required upgrades to keep the system operational at an agreed service level. Maintainer could also be charged with abortion of maintenance and disposal of the system.
- *Regulator (R)*: The entity that enforces standards and laws for the protection of life, property, or the natural habitat through imposing duties and accreditation/certification.

6.4 The levels of ethical privacy certification

Three main levels of assessment of conformity are established, depending on the scale of risks posed and the impact of the AIS on health, welfare, safety, and ethical values of stakeholders. The levels are:

- *Baseline, low impact (LI)*: The smallest subset of privacy EFRs is applicable for conformity assessment.
- *Compliant, medium impact (MI)*: A larger set of privacy EFRs than baseline is applicable for conformity assessment.
- *Critical, high impact (HI)*: Any AIS product, service, or system that presents a likelihood of injury or harm to well-being, health, safety, security, and welfare must satisfy all ethical privacy EFRs.

The level of certification is determined through a risk-profiling exercise on the product, service, or system that takes place as the first phase of the conformity assessment activities.

6.5 Required evidence

These are the types and quantity of evidence items required to satisfy the stated requirements. A single requirement may relate to one or many items of objective evidence for evaluation of the degree to which the requirement is met (satisfaction).

6.6 Evaluation of evidence

The evaluation of evidence comprises a suitable scale of measurement and scoring of the evidence. A two-tier approach to the measurement of the evidence items is adopted as follows:

- a) Top-level finding: No critical findings in the detailed normative requirements/areas requiring attention for improvement.
- b) Overall score: On a 1 to 5 scale (based on aggregate of satisfying sublevel goals):
 - 5- Excels baseline requirements
 - 4- Sustains baseline requirements
 - 3- Meets baseline requirements (typical pass mark)
 - 2- Needs improvement
 - 1- Does not meet requirements

A score of 3 is generally considered to be a sufficient pass mark for most cases. However, certain elements that represent a particularly strong risk or that operate in a mission-critical capacity may require a higher score to be considered sufficient.

NOTE 1—The scale of evaluation and the typical pass mark shall be appropriate to the criticality of the requirement and the nature of the evidence and may vary for each privacy EFR.

NOTE 2—Each privacy EFR can have its own bespoke units, measurement scale, and benchmark for evaluation appropriate to its nature. The 1 to 5 scoring adopted is the default for all privacy EFRs in Annex B and can be modified as appropriate to the nature of the evidence.

6.7 The constraints of ethical privacy certification

The certification process cannot cover every potential eventuality. Changes in technology, culture, law, consumer standards, and practices may diminish its effectiveness or applicability to support the quality of ethical privacy. Eventually, without update, the certification may drift from contemporary realities and established best practices.

Therefore, it will be important to make regular updates and amendments to the underlying concept schema where appropriate. The IEEE CertifAIEd™ team has forecast potential technological and cultural developments for a foreseeable time horizon, thereby future proofing the criteria and certification as far as possible. This has been accomplished through discussion of technologies or practices that may be prototyped presently but are not yet in common deployment or in line with established norms and best practices.

Annex A

AIS ethical privacy schema

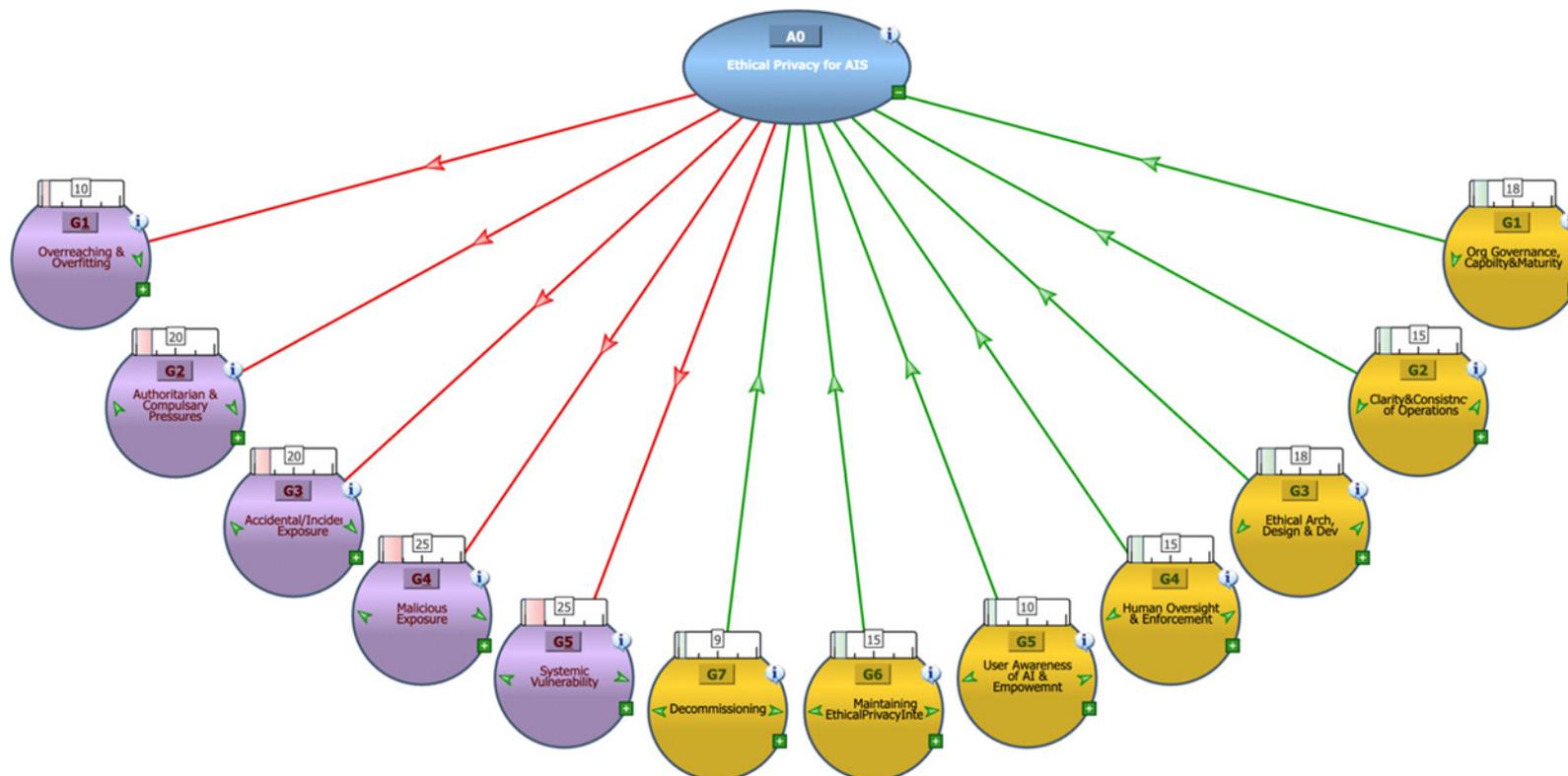


Figure A.1—Drivers and inhibitors of AIS ethical privacy.

Annex B

Ethical privacy certification criteria

Privacy schema goal description	Privacy ethical foundational requirements (EFRs)	Normative/instructive	Cert level LI, MI, HI	Duty holder D, I, O, M, R	Required evidence	Evidence measurement and typical pass mark
<p>G1 - Organizational governance, capability, and maturity</p> <p>The organization’s capability, maturity, governance processes, and political will/good faith for ethical privacy assurance</p>	<p>The following privacy ethical foundational requirements shall be fulfilled for the product, system, or service by the duty holders:</p> <p>a) The organization shall have in place a governance and oversight framework that puts ethical privacy into practice and can provide assurance of capable and mature adherence across the organization and across the AIS life cycle.</p>	N	MI	D, I, O, M, R	<p>The following items of evidence fulfill the foundational requirements.</p> <p>a) Certification that the organization adheres to some other quality standard in relation to its governance or compliance procedures</p> <p>b) A copy of an operations manual including coordinated management and monitoring of ethical privacy across all roles and operational contexts</p> <p>c) A copy of the organizational chart highlighting designated lines of responsibility, accountability, consultation, and information (RACI) flows within the organization particularly concerning ethical privacy</p> <p>d) Details of engagement and participation in industry/regulatory initiatives concerning ethical privacy</p>	<p>Two-tier approach measurement of the evidence items:</p> <p>a) Top-level finding: “No critical findings in the detailed normative requirements”/“areas requiring attention for improvement.”</p> <p>b) Overall score: On 1-5 scale (based on aggregate of satisfying sublevel goals) such as:</p> <p>5- Excels baseline requirements 4- Sustains baseline requirements 3- <u>Meets baseline requirements</u> (typical pass mark) 2- Needs improvement 1- Does not meet requirements</p>
<p>G2 - Clarity and consistency of operations of AIS</p> <p>Clear definition, articulation, and communication of the</p>	<p>The following privacy ethical foundational requirements shall be fulfilled for the product, system, or service by relevant duty holders:</p>				<p>The following items of evidence are required in fulfillment of the foundational requirements:</p> <p>a) Diagrammatic and textual records of ConOps for the product,</p>	<p>Two-tier approach measurement of the evidence items:</p> <p>a) Top-level finding: “No critical findings in the detailed normative</p>

Privacy schema goal description	Privacy ethical foundational requirements (EFRs)	Normative/instructive	Cert level LI, MI, HI	Duty holder D, I, O, M, R	Required evidence	Evidence measurement and typical pass mark
<p>concepts and results of operation in the intended environments for the AIS product, service, or system to the relevant stakeholders</p>	<p>a) Provide a clear definition, articulation, and communication of the concepts and results of operation in the intended environments for the AIS product, service, or system to the relevant stakeholders</p>	<p>N</p>	<p>HI</p>	<p>D, I, O</p>	<p>service, or system</p> <p>b) Records indicating consideration of ConOps within all intended environments of application</p> <p>c) Records indicating consideration of key stakeholders in all environments of application</p> <p>d) Records indicating communication of the ConOps to all stakeholders in a suitable and appropriate language for each group</p>	<p>requirements”/“areas requiring attention for improvement.”</p> <p>b) Overall score: On 1-5 scale (based on aggregate of satisfying sublevel goals) such as:</p> <p>5- Excels baseline requirements</p> <p>4- Sustains baseline requirements</p> <p>3- <u>Meets baseline requirements</u> (typical pass mark)</p> <p>2- Needs improvement</p> <p>1- Does not meet requirements</p>
<p>G3 - Ethical architecture, design and development for AIS</p> <p>Upholding a holistic approach to ethical design and development at all levels of the organization, empowering staff to review the activities and focus of peers and provide feedback; due regard to all attributes and aspects of the architecture, design, and development that could be invasive to ethical privacy (e.g., holistic, consultative, and provision for feedback</p>	<p>The following privacy ethical foundational requirements shall be fulfilled for the product, system, or service by relevant duty holders:</p> <p>a) Responsible parties should promote a culture of peer accountability and opportunities to raise concerns and hold space for discussion within the organization</p> <p>b) Provide opportunities and interfaces for stakeholders who have indirect influence but</p>	<p>N</p> <p>N</p>	<p>HI</p> <p>MI</p>	<p>D, I, O, M, R</p> <p>D, I, O, M, R</p>	<p>The following items of evidence are required in fulfillment of the foundational requirements:</p> <p>a) Evidence of a change management strategy/plan that promotes a culture of engagement and active discussion of issues and solutions</p> <p>b) Evidence of mechanisms in place that facilitate frank and honest discussion of ethical and safety concerns within the organization</p> <p>c) Absence of procedural constraints or incentives that could be likely to frustrate such processes</p> <p>d) Provision of opportunities for</p>	<p>Two-tier approach measurement of the evidence items:</p> <p>a) Top-level finding: “No critical findings in the detailed normative requirements”/“areas requiring attention for improvement.”</p> <p>b) Overall score: On 1-5 scale (based on aggregate of satisfying sublevel goals) such as:</p> <p>5- Excels baseline requirements</p> <p>4- Sustains baseline</p>

Privacy schema goal description	Privacy ethical foundational requirements (EFRs)	Normative/instructive	Cert level LI, MI, HI	Duty holder D, I, O, M, R	Required evidence	Evidence measurement and typical pass mark
and fulfillment of the ethical privacy requirements)	<p>no responsibility nor accountability to raise concerns and discuss issues and solutions</p> <p>c) At all stages of development, from initial design preproduction through to final decommissioning, products must be designed with ethical regard to privacy requirements</p>	N	HI	D, I, O, M, R	<p>stakeholders to understand systems and procedures in a clear yet comprehensive manner</p> <p>e) Provision of opportunities to provide feedback and procedures in place to collect such information and apply it to potential design revisions; such information is to be easily accessible within a knowledge management system that tracks the outcomes to ensure cross referencing and reuse</p>	<p>requirements</p> <p>3- <u>Meets baseline requirements</u> (typical pass mark)</p> <p>2- Needs improvement</p> <p>1- Does not meet requirements</p>
<p>G4 - Human oversight and enforcement in AIS</p> <p>Human agents should be able to understand the AIS product, service, or system behavior in order to be able to intervene and set up a process to deny continuation of activity and assess the context to ensure timely corrective action; in implementing human oversight, the organization should be mindful of and mitigate against harmful or detrimental intervention including the risks to ethical privacy due to human oversight</p>	<p>The following privacy ethical foundational requirements shall be fulfilled for the product, system, or service by relevant duty holders:</p> <p>a) The organization shall ensure that at all stages of the AIS life cycle the AIS product, service, or system behaviors and outcomes are clear and understandable to those seeking to govern and oversee its functions.</p> <p>b) The organization shall have processes and procedures in place to continually monitor the AIS system’s intended goals against actual outcomes for ethical privacy.</p>	N	MI	D, I, O, M, R	<p>The following items of evidence are required in fulfillment of the foundational requirements:</p> <p>a) Records and documentation used internally to explain clearly and unequivocally to the ethical privacy governance and oversight resource the intended goal of the AIS product, service, or system, how it is intended to operate, and how and why it achieves the outcomes</p> <p>b) Process and procedures for continuous monitoring of AIS intended goals against actual outcomes to ensure ethical privacy is not being eroded or invaded</p> <p>c) A copy of the risk management policy detailing intervention (including fail-safe instructions) and corrective actions, including</p>	<p>Two-tier approach measurement of the evidence items:</p> <p>a) Top-level finding: “No critical findings in the detailed normative requirements”/“areas requiring attention for improvement.”</p> <p>b) Overall score: On 1-5 scale (based on aggregate of satisfying sublevel goals) such as:</p> <p>5- Excels baseline requirements</p> <p>4- Sustains baseline requirements</p> <p>3- <u>Meets baseline requirements</u> (typical pass mark)</p> <p>2- Needs improvement</p>

Privacy schema goal description	Privacy ethical foundational requirements (EFRs)	Normative/instructive	Cert level LI, MI, HI	Duty holder D, I, O, M, R	Required evidence	Evidence measurement and typical pass mark
	<ul style="list-style-type: none"> c) The organization shall have in place a risk management policy describing intervention and corrective action in the event that either the whole or an element of the AIS system fails to adhere to the organization’s ethical privacy principles. d) The organization shall ensure the human oversight in place is meaningful concerning ethical privacy. 				<ul style="list-style-type: none"> timings, to be taken in the event of an AIS system risk being identified as eroding or invading ethical privacy d) Details of how the human oversight concerning ethical privacy has real decision-making power, maintains its independence, and is apprised not to reintroduce bias into any AIS system which could impact ethical privacy 	<p>1- Does not meet requirements</p>
<p>G5 – End-user awareness of AIS and empowerment</p> <p>Potential users being aware of the existence and functions of an AI element within the product, service, or system in the context of use and being empowered to sufficiently understand and make decisions of the use of such systems; may disadvantage the end user due to lack of suitable alternative options</p>	<p>The following privacy ethical foundational requirements shall be fulfilled for the product, system, or service by relevant duty holders:</p> <ul style="list-style-type: none"> a) Ensure availability and transparency to end user of lay-person information about the AIS to enable accurate evidence-based decisions 	<p>N</p>	<p>H</p>	<p>D, I, O</p>	<p>The following items of evidence are required in fulfillment of the foundational requirements:</p> <ul style="list-style-type: none"> a) Accessibility of information required for end user to make informed decisions, including access by sight-impaired b) Mechanism for human contact to clarify information and/or receive additional information not located by end user (within parameters of competition constraints) 	<p>Two-tier approach measurement of the evidence items:</p> <ul style="list-style-type: none"> a) Top-level finding: “No critical findings in the detailed normative requirements”/“areas requiring attention for improvement.” b) Overall score: On 1-5 scale (based on aggregate of satisfying sublevel goals) such as: <ul style="list-style-type: none"> 5- Excels baseline requirements 4- Sustains baseline requirements 3- <u>Meets baseline requirements</u>

Privacy schema goal description	Privacy ethical foundational requirements (EFRs)	Normative/instructive	Cert level LI, MI, HI	Duty holder D, I, O, M, R	Required evidence	Evidence measurement and typical pass mark
						<p><u>(typical pass mark)</u> 2- Needs improvement 1- Does not meet requirements</p>
<p>G6 - Maintaining ethical privacy integrity</p> <p>Efforts to maintain an ethical profile of the AIS product, service, or system with respect to privacy requirements and criteria/behaviors</p>	<p>The following privacy ethical foundational requirements shall be fulfilled for the product, system, or service by relevant duty holders:</p> <p>a) Maintain an ethical profile of the AIS product, service, or system with respect to privacy requirements and criteria/behaviors</p>	<p>N</p>	<p>LI</p>	<p>D, I, O, M</p>	<p>The following items of evidence are required in fulfilment of the foundational requirements:</p> <p>a) Records demonstrating awareness of key privacy aspects pertaining to the AIS product, service, or system</p> <p>b) Records indicating that positive policies, processes, and procedures are devised and actions taken to incorporate privacy control features into the AIS</p> <p>c) Records indicating that the privacy-related features and profile for the AIS have been monitored and maintained at stakeholder acceptable level over the life cycle</p>	<p>Two-tier approach measurement of the evidence items:</p> <p>a) Top-level finding: “No critical findings in the detailed normative requirements”/“areas requiring attention for improvement.”</p> <p>b) Overall score: on 1-5 scale (based on aggregate of satisfying sublevel goals) such as:</p> <p>5- Excels baseline requirements 4- Sustains baseline requirements 3- <u>Meets baseline requirements</u> <u>(typical pass mark)</u> 2- Needs improvement 1- Does not meet requirements</p>
<p>G7 - Decommissioning</p> <p>Risk and control mechanisms put in place in the decommissioning of an AIS concerning data (anonymization/deletion),</p>	<p>The following ethical privacy foundational requirements shall be fulfilled for the product, system, or service by relevant duty holders:</p> <p>a) Entities designing,</p>	<p>N</p>	<p>MI</p>	<p>I. D, I, O, M, R II. O, M III. O, M IV. O, M</p>	<p>The following items of evidence are required in fulfillment of the foundational requirements:</p> <p>a) Entities involved in the design, development, sale, or other provision of any AIS shall</p>	<p>Two-tier approach measurement of the evidence items:</p> <p>a) Top-level finding: “No critical findings in the detailed normative requirements”/“areas</p>

Privacy schema goal description	Privacy ethical foundational requirements (EFRs)	Normative/instructive	Cert level LI, MI, HI	Duty holder D, I, O, M, R	Required evidence	Evidence measurement and typical pass mark
<p>metadata, insight and inference, learning, and legacy code or models.</p>	<p>developing, selling, providing, and otherwise making available AIS products or services shall provide operators of AIS systems with step-by-step instructions for decommissioning an AIS based on their respective knowledge/ understanding of potential harm and unintended consequences associated with the inputs and outputs of such systems during development and operation.</p> <p>For the purposes of this section (G7), decommissioning shall include the turning off and removal of the AIS system and all of its components.</p> <p>b) Entity controlling the AIS at the time of decommissioning shall conform to all local laws and contractual requirements concerning the data associated with the AIS, with particular attention to data that is</p>				<p>include in their instruction and training materials, and in their maintenance or service agreements (where applicable), instructions and advice relating to effective decommissioning of such AIS systems.</p> <p>b) Entity (or entities) controlling AIS at time of decommissioning shall have in place a policy that reflects common best practices for AIS decommissioning within their industry, sector, or jurisdiction as appropriate.</p> <p>c) Entities controlling AIS at the time of decommissioning shall maintain logs indicating the steps taken to decommission the AIS, including the actions called for by the AIS designer and developer materials.</p>	<p>requiring attention for improvement.”</p> <p>b) Overall score: On 1-5 scale (based on aggregate of satisfying sublevel goals) such as:</p> <p>5- Excels baseline requirements 4- Sustains baseline requirements 3- <u>Meets baseline requirements</u> (<u>typical pass mark</u>) 2- Needs improvement 1- Does not meet requirements</p>

Privacy schema goal description	Privacy ethical foundational requirements (EFRs)	Normative/instructive	Cert level LI, MI, HI	Duty holder D, I, O, M, R	Required evidence	Evidence measurement and typical pass mark
	<p>defined as personal information, personal identifiable information (PII), personal data, and so forth, in any of the jurisdictions in which the entity operates or has end users/customers.</p> <p>c) Entity controlling the AIS operation has a duty to assure that all data inputs (and data outputs) of the AIS associated with all stages of its life cycle are deleted.</p>					
<p>G1b - Overreaching and overfitting</p> <p>Use of technologies that overstep the bounds of dignity or appropriateness by either overfitting of certain characteristics or drawing unreasonable inferences based upon isolated data points; this could include unwarranted and unexpected (from the user’s perspective) cross correlation of data sets</p>	<p>The following privacy ethical foundational requirements shall be fulfilled for the product, system, or service by relevant duty holders:</p> <p>a) AIS stakeholders shall create and maintain policies and communications materials and programs in a continuous effort to keep human data subjects of input data and output data apprised and informed about the current and anticipated inference capabilities of their systems.</p>	N	LI	O, M, R	<p>The following items of evidence are required in fulfillment of the foundational requirements:</p> <ul style="list-style-type: none"> a) Copies of policies and communications materials b) Reports of market research involving data subjects’ expectations of AIS capabilities 	<p>Two-tier approach measurement of the evidence items:</p> <ul style="list-style-type: none"> a) Top-level finding: “No critical findings in the detailed normative requirements”/“areas requiring attention for improvement.” b) Overall score: On 1-5 scale (based on aggregate of satisfying sublevel goals) such as: <ul style="list-style-type: none"> 5- Excels baseline requirements 4- Sustains baseline requirements 3- <u>Meets baseline</u>

Privacy schema goal description	Privacy ethical foundational requirements (EFRs)	Normative/instructive	Cert level LI, MI, HI	Duty holder D, I, O, M, R	Required evidence	Evidence measurement and typical pass mark
	b) AIS stakeholders shall maintain a steady market research activity to enable them to identify and close gaps between data subject understanding and reality of AIS capabilities.					<u>requirements (typical pass mark)</u> 2- Needs improvement 1- Does not meet requirements
G2b - Authoritarian and compulsory pressures Demands and the ability by some institutions or governmental bodies to gain access to information on AIS as held by a given organization/duty holder	The following privacy ethical foundational requirements shall be fulfilled for the product, system, or service by relevant duty holders: a) Each AIS stakeholder will maintain a publicly available policy that clearly indicates its processes and policies relating to governmental and law enforcement activities relating to that AIS system. b) Where an AIS stakeholder is not permitted to reveal the power or action of a governmental authority to intrude on an AIS system, it shall be transparent in so far as it is lawfully and safely able to do so. c) Should a vendor opt to sell AIS into a regime	N	LI	D, I, O, M, R	The following items of evidence are required in fulfillment of the foundational requirements: a) Policy documents as posted by AIS stakeholder	Two-tier approach measurement of the evidence items: a) Top-level finding: “No critical findings in the detailed normative requirements”/“areas requiring attention for improvement.” b) Overall score: On 1-5 scale (based on aggregate of satisfying sublevel goals) such as: 5- Excels baseline requirements 4- Sustains baseline requirements 3- <u>Meets baseline requirements (typical pass mark)</u> 2- Needs improvement 1- Does not meet requirements

Privacy schema goal description	Privacy ethical foundational requirements (EFRs)	Normative/instructive	Cert level LI, MI, HI	Duty holder D, I, O, M, R	Required evidence	Evidence measurement and typical pass mark
	<p>that is not in alignment with the expectations of those who provided their data, a new system trained on approved user data must be provided to maintain privacy.</p>					
<p>G3b - Accidental/incidental exposure</p> <p>Inadvertent loss or breach of security and loss of control of AIS (data, the system, and the platform) including eavesdropping and acquisition or interception of downstream data likely to compromise.</p>	<p>The following privacy ethical foundational requirements shall be fulfilled for the product, system, or service by relevant duty holders:</p> <ul style="list-style-type: none"> a) Take appropriate measures to avoid inadvertent loss or breach of security and loss of control of AIS (data, the system, and the platform) including eavesdropping and acquisition or interception of downstream data likely to compromise privacy 	N	MI	D, I, O, M, R	<p>The following items of evidence are required in fulfillment of the foundational requirements:</p> <ul style="list-style-type: none"> a) Policies and procedures indicating proactive risk assessment and mitigation measures for avoidance of accidental or incidental exposure of data likely to compromise privacy of stakeholders b) An incident log detailing any data exposure and remedial actions taken c) Records indicating communication of data exposure with the relevant stakeholders 	<p>Two-tier approach measurement of the evidence items:</p> <ul style="list-style-type: none"> a) Top-level finding: “No critical findings in the detailed normative requirements”/“areas requiring attention for improvement.” b) Overall score: On 1-5 scale (based on aggregate of satisfying sublevel goals) such as: <ul style="list-style-type: none"> 5- Excels baseline requirements 4- Sustains baseline requirements 3- <u>Meets baseline requirements (typical pass mark)</u> 2- Needs improvement 1- Does not meet requirements
<p>G4b - Malicious exposure</p> <p>Intentional breach of security and loss of control of</p>	<p>The following privacy ethical foundational requirements shall be fulfilled for the product, system, or service by relevant</p>				<p>The following items of evidence are required in fulfillment of the foundational requirements:</p>	<p>Two-tier approach measurement of the evidence items:</p> <ul style="list-style-type: none"> a) Top-level finding: “No

Privacy schema goal description	Privacy ethical foundational requirements (EFRs)	Normative/instructive	Cert level LI, MI, HI	Duty holder D, I, O, M, R	Required evidence	Evidence measurement and typical pass mark
system/data or unauthorized access to the data/system including eavesdropping and acquisition or interception of downstream data likely to compromise privacy.	duty holders:					
	a) Responsible parties shall take care to prevent the exposure of private details of stakeholders to whom they have a duty of care	N	HI	D, I, O, M, R	a) Evidence of adequate security measures including, but not necessarily limited to, cybersecurity and operational security according to industry best practices, standards, and certifications	critical findings in the detailed normative requirements?/"areas requiring attention for improvement."
	b) Necessitated investment in adequate security mechanisms, including protection against potential lapses by operators	N	HI	D, I, O, M, R	b) Secured access to any monitoring and administration suites, including the logging of access to what details, by whom, from which location, device, MAC address, IP, and so forth, and for what purpose	b) Overall score: On 1-5 scale (based on aggregate of satisfying sublevel goals) such as:
	c) Protection to prevent monitoring systems intended for administration and research purposes from being abused for eavesdropping or other unauthorized purposes	N	HI	D, I, O, M	c) Engaging only with data partners or other parties whose past and present conduct is credible and inspires trust and confidence	5- Excels baseline requirements 4- Sustains baseline requirements 3- <u>Meets baseline requirements</u> (typical pass mark)
	d) Taking care to pass data only on to parties that one has confidence in their security mechanisms and to secure such flows of data in a secure manner also to prevent interception	N	HI	D, I, O, M, R	d) Adequate encryption for data flows internal and external to the organization.	2- Needs improvement 1- Does not meet requirements
e) Appropriate and qualified resourcing to oversee and ensure the functioning of the mechanisms to mitigate malicious activity				e) Organizational design of human oversight		

Privacy schema goal description	Privacy ethical foundational requirements (EFRs)	Normative/instructive	Cert level LI, MI, HI	Duty holder D, I, O, M, R	Required evidence	Evidence measurement and typical pass mark
<p>G5b - Systemic vulnerability</p> <p>Structural stochasticity in the AIS system that can pose a risk to/undermine privacy. This may appear in different components and may not be a permanent state of AIS. Any time an algorithm is transferred from one system to other, this phenomenon may occur.</p>	<p>The following privacy ethical foundational requirements shall be fulfilled for the product, system, or service by relevant duty holders:</p> <ul style="list-style-type: none"> a) Responsible parties shall provide adequate forethought and hardening against potential built-in elements of randomness that can create error machine learning (ML) systems may harbor, whether due to a semi-stochastic nature of an ML technique or the variations in data, or differences between lab conditions and real life. 	<p>N</p>	<p>HI</p>	<p>D, I, O, M, R</p>	<p>The following items of evidence are required in fulfillment of the foundational requirements:</p> <ul style="list-style-type: none"> a) System design documents that specify validation and verification methods and protocols employed to reduce the likelihood and impact of systemic vulnerability due to random events and variations; such methods may include (but are not limited to) ensemble models, hyperparameter tuning, reference to other data sources, system audits, and verification by human intelligence b) Organizational design detailing the resources responsibly competent to oversee the validation and verification protocols to mitigate structural stochasticity 	<p>Two-tier approach measurement of the evidence items:</p> <ul style="list-style-type: none"> a) Top-level finding: “No critical findings in the detailed normative requirements”/“areas requiring attention for improvement.” b) Overall score: On 1-5 scale (based on aggregate of satisfying sublevel goals) such as: <ul style="list-style-type: none"> 5- Excels baseline requirements 4- Sustains baseline requirements 3- <u>Meets baseline requirements</u> (<u>typical pass mark</u>) 2- Needs improvement 1- Does not meet requirements
<p>END</p>						

Annex C

Bibliography

The following sources and public domain frameworks have been consulted for the verification, coverage, integrity, quality, and currency of the certification criteria independently developed in IEEE CertifAIEd™.

[B1] “Ethics Guidelines for Trustworthy AI,” High-Level Expert Group on Artificial Intelligence (AI HLEG), European Commission, Apr. 2019.⁴

[B2] *The Age of Digital Interdependence*, Report of the UN Secretary-General’s High-level Panel on Digital Cooperation, United Nations, Jun. 2019.⁵

[B3] OECD/LEGAL/0449, *Recommendation of the Council on Artificial Intelligence*, May 21, 2019.⁶

[B4] 2019 “G20 AI Principles,” *G20 Ministerial Statement on Trade and Digital Economy*, Annex, Jun. 2019.⁷

[B5] “Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems,” The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, Apr. 4, 2019.

[B6] Floridi, L., J. Cowls, T. C. King, “How to design AI for social good: Seven essential factors,” *Science and Engineering Ethics*, vol. 26, pp.1771–1796, 2020.

[B7] Madary, M., and Thomas K. Metzinger, “Real virtuality: A code of ethical conduct. recommendations for good scientific practice and the consumers of VR-technology,” *Frontiers in Robotics and AI*, vol. 3, no. 3, Feb. 19, 2016.

[B8] “The State of AI Ethics,” Montreal AI Ethics Institute, Jan. 2021.⁸

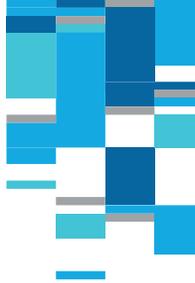
⁴ European Commission publications are available from the Futurium website (<https://futurium.ec.europa.eu/en>).

⁵ United Nations publications are available from the United Nations website (<https://www.un.org/>).

⁶ Organisation for Economic Co-operation and Development publications available from OECD Legal Instruments website (<https://legalinstruments.oecd.org/>).

⁷ Available from: <https://www.mofa.go.jp/files/000486596.pdf>.

⁸ Available from <https://montrealaiethics.ai/wp-content/uploads/2021/01/The-State-of-AI-Ethics-Report-January-2021.pdf>.



IEEE CertifAIEd™

<http://engagestandards.ieee.org/ieeecertifaiied.html>

Connect with us on:

-  **Twitter:** twitter.com/ieeesa
-  **Facebook:** facebook.com/ieeesa
-  **LinkedIn:** linkedin.com/groups/1791118
-  **Beyond Standards blog:** beyondstandards.ieee.org
-  **YouTube:** youtube.com/ieeesa

standards.ieee.org
Phone: +1 732 981 0060