

I. PREFACE

Dear entrepreneurs, dear organizations!

The present “Trustworthy AI – by Know Center” audit guideline (“guideline”) sets out the award criteria for the Know Center certification “Trustworthy AI – by Know Center.” The certification has been developed based on recognized principles of trustworthy artificial intelligence, expanded to include key aspects of digital sovereignty, and supplemented by insights from research and practice.

By applying for the “Trustworthy AI – by Know Center” certification, you opt for a quality management system of above-average value in the field of trustworthy AI. This achievement is communicated through the “Trustworthy AI – by Know Center” certification.

The guideline is based on the following principles:

- Data protection and security
- Transparency, explainability, and fairness
- Human oversight and decision-making
- Digital sovereignty
- Economic viability
- Sustainability
- Digital humanism

Applying for the “Trustworthy AI – by Know Center” certification is voluntary and implies consent to undergo an assessment based on the criteria set out in this guideline.

The certification contributes to strengthening trust in AI, enhancing digital sovereignty, and supporting Austria’s/Europe’s position as a business and innovation hub.

Kind regards

Dr. S. Cathrin von Hesler

Site Manager Wien



©Natascha Unkart

Dr. Kerstin Waxnegger

*Senior Legal Counsel & Data
Protection Officer*



©Openfoto Hausleitner

II. GUIDING PRINCIPLES

The quality assurance system underpinning the certification focuses on the areas of trustworthy AI and digital sovereignty.

Trustworthy AI refers to artificial intelligence systems that are designed, developed, and operated in accordance with ethical, legal, and technical standards, ensuring that they behave reliably and in a comprehensible manner.

Key aspects include human oversight, transparency, fairness, robustness, and data protection. Trustworthy AI requires an interdisciplinary approach that combines technological innovation with ethical guidelines and regulatory requirements. Only through the consistent integration of these principles can societal trust in AI systems be sustainably strengthened.

The development and deployment of trustworthy AI make a significant contribution to digital sovereignty. Digital sovereignty describes the ability of organisations and states to develop, operate, and govern digital technologies in a self-determined, secure, and legally compliant manner. Trustworthy AI strengthens this capability by ensuring transparency, traceability, robustness, fairness, and data protection, thereby enabling the controlled use and independent assessment of AI systems and models. It forms part of a broader understanding of digital sovereignty, which also encompasses technological infrastructure, data availability, interoperability, and organisational capabilities.

Consistent adherence to the principles of trustworthy AI strengthens societal trust and enables the controlled and sovereign governance of digital infrastructures, thereby supporting autonomy in key sectors such as healthcare, mobility, energy supply, and industry.

Trustworthy AI and digital sovereignty together form the foundation for the sustainable and self-determined shaping of digital transformation at both national and European levels.

III. GENERAL AWARDING PROVISIONS

1. Definitions

The terms *quality seal* and *certification* are used synonymously.

2. Voluntary participation

The application for the "Trustworthy AI – by Know Center" certification seal is voluntary and implies consent to undergo the assessment criteria set out in the guideline.

3. Licence agreement

As a prerequisite for the use of the "Trustworthy AI – by Know Center" certification for purposes of labelling, marketing, or advertising, the conclusion of a valid licence agreement with Know Center is required. Upon conclusion of the licence agreement, the right to use the "Trustworthy AI – by Know Center" certification is granted.

Licensees are all natural or legal persons who have entered into an agreement with Know Center for the use of the "Trustworthy AI – by Know Center" certification. The licence agreement grants the right to use the protected "Trustworthy AI – by Know Center" certification.

IV. SPECIAL PROCUREMENT PROVISIONS

Two provisions are defined as specific criteria of the procurement procedure.

The first provision concerns the assessment conducted by the Know Center during the **research and/or methods and prototype development phase**.

The second provision concerns the applicant's self-assessment procedure and the applicant's obligation to confirm that such a self-assessment has been carried out.

1. First procurement provision

The first prerequisite for award is the Trustworthy AI assessment conducted during the **research and/or prototype phase** with respect to the following criteria: data protection and security, transparency, explainability and fairness, human oversight and decision-making, digital sovereignty, economics, sustainability, and digital humanism.

The "Trustworthy AI – by Know Center" quality label for artificial intelligence is based on the following criteria:

- Data protection and security

Technical security: The primary focus is operational safety; it must be ensured that no individuals are harmed.

Functional safety: The AI system must behave in a predictable and robust manner in safety-critical systems.

Data security and data protection: The protection of personal data and other sensitive information during training and operation is a key priority.

- Transparency, Explainability and Fairness

Traceability of outputs is supported through application-specific selection of technical methods.

A sound understanding of bias and fairness in both data and application contexts forms the basis for equitable and non-discriminatory outcomes.

- Human Oversight and Decision-Making

The preservation of human autonomy is central.

In addition, a task-appropriate division of responsibilities between humans and AI is to be developed.

- Digital Sovereignty

Data sovereignty: The processing of personal and other protected data must be appropriately safeguarded.

The research location is within Europe, and the research services are provided by a European company.

Furthermore, dependence on individual large technology corporations is to be avoided.

- Economics

Increased autonomy of processes leads to higher productivity.

- Sustainability

Energy consumption: Given the high energy requirements for training large AI models, emphasis must be placed on energy-efficient algorithms and hardware.

Sustainable solutions are achieved through the involvement of relevant stakeholders, enabling structural problem-solving via AI.

- Digital Humanism

The human being must remain at the center; AI is to serve people – not technology.

Furthermore, the application must be aligned with human values in accordance with its use context.

Abuse prevention: Harmful use of the AI system shall be actively countered.

2. Second procurement provision

Following the completion of the initial assessment pursuant to Section V.1 by Know Center, a second prerequisite for award is the applicant's self-monitoring, to be carried out at defined intervals. The applicant must confirm to Know Center, acting as the certifying body, that such self-monitoring is being conducted.

For this purpose, the applicant must develop a self-monitoring concept based on the criteria set out under V.1. Furthermore, the applicant must define the frequency of self-monitoring. The concept must also describe how AI competency is ensured. No further explanation regarding AI competency is required if the applicant can provide evidence of a relevant competency assessment of its employees according to the DigComp standard (Digital Competence Framework EU, AT or equivalent). The applicant must submit the concept to Know Center as the certifying body.

The results of the self-monitoring must be documented by the applicant. Documentation of the self-assessment must be retained for a period of five years.

V. AWARDING OF THE CERTIFICATION

Upon fulfilment of the criteria, Know Center grants the certification and notifies the company in writing of the assigned certification ID. In addition, the applicant receives a certificate confirming the right to use the certification, as well as the corresponding seal for further use. The applicant may only access the seal via their IP address (privacy by design).

As the qualifying institution, Know Center retains discretionary authority in the evaluation process. There is no legal entitlement to the awarding of the certification or the certification ID.

VI. VALIDITY

The validity of the certification is 24 months. An extension is possible provided that the applicable standards continue to be met. An application for renewal may be rejected by Know Center if amended awarding criteria apply, for example due to changes in legal requirements.

VII. REVOCATION

The certification may be revoked if the standards defined in this guideline are no longer met. Know Center shall have discretionary authority in assessing compliance. The decision shall be communicated in writing. Upon revocation, the right to use the "Trustworthy AI – by Know Center" certification shall expire.

Fees are non-refundable.

APPLICABILITY

This guideline applies as of April 1st 2026

LEGAL NOTICE

Media owner and publisher: Know Center Research GmbH

A-8010 Graz, Sandgasse 34, Tel: +43 316 873 30801, info@know-center.at

Design: corporate identity prihoda gmbH

© 2026

Copying and distribution permitted only in unaltered form!