

ETSI TR 104 159 V1.1.1 (2026-01)



TECHNICAL REPORT

**Securing Artificial Intelligence (SAI);
Understanding and Preventing Harm from Generative AI**
(<https://standards.iteh.ai>)
Document Preview

[ETSI TR 104 159 V1.1.1 \(2026-01\)](https://standards.iteh.ai)

<https://standards.iteh.ai/catalog/standards/etsi/a82bc5c7-e436-4120-96e4-d8447e674da8/etsi-tr-104-159-v1-1-1-2026-01>

Reference

DTR/SAI-0019

Keywords

AI, cybersecurity, end-user, privacy

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - APE 7112B
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° w061004871

Important notice

The present document can be downloaded from the
[ETSI Search & Browse Standards](#) application.

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the prevailing version of an ETSI deliverable is the one made publicly available in PDF format on [ETSI deliver](#) repository.

Users should be aware that the present document may be revised or have its status changed, this information is available in the [Milestones listing](#).

If you find errors in the present document, please send your comments to the relevant service listed under [Committee Support Staff](#).

If you find a security vulnerability in the present document, please report it through our [Coordinated Vulnerability Disclosure \(CVD\)](#) program.

Notice of disclaimer & limitation of liability

The information provided in the present deliverable is directed solely to professionals who have the appropriate degree of experience to understand and interpret its content in accordance with generally accepted engineering or other professional standard and applicable regulations.

No recommendation as to products and services or vendors is made or should be implied.

No representation or warranty is made that this deliverable is technically accurate or sufficient or conforms to any law and/or governmental rule and/or regulation and further, no representation or warranty is made of merchantability or fitness for any particular purpose or against infringement of intellectual property rights.

In no event shall ETSI be held liable for loss of profits or any other incidental or consequential damages.

Any software contained in this deliverable is provided "AS IS" with no warranties, express or implied, including but not limited to, the warranties of merchantability, fitness for a particular purpose and non-infringement of intellectual property rights and ETSI shall not be held liable in any event for any damages whatsoever (including, without limitation, damages for loss of profits, business interruption, loss of information, or any other pecuniary loss) arising out of or related to the use of or inability to use the software.

Copyright Notification

No part may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm except as authorized by written permission of ETSI.

The content of the PDF version shall not be modified without the written authorization of ETSI.

The copyright and the foregoing restriction extend to reproduction in all media.

© ETSI 2026.
All rights reserved.

Contents

Intellectual Property Rights	5
Foreword.....	5
Modal verbs terminology.....	5
1 Scope	6
2 References	6
2.1 Normative references	6
2.2 Informative references.....	6
3 Definition of terms, symbols and abbreviations.....	8
3.1 Terms.....	8
3.2 Symbols.....	9
3.3 Abbreviations	9
4 Introduction	10
4.1 What is Generative Artificial Intelligence (GenAI)	10
4.2 Uses of GenAI.....	10
4.3 Impact of Regulation and Legislation in a Global Perspective.....	10
4.3.1 Introduction.....	10
4.3.2 Australia - Voluntary AI Safety Standard.....	10
4.3.3 Brazilian Legal Framework for Artificial Intelligence, Marco Legal da Inteligência Artificial (Bill No. 2338/2023).....	11
4.3.4 Canada - Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems	11
4.3.5 China.....	12
4.3.5.1 The Interim Measures for the Management of Generative AI Services.....	12
4.3.5.2 Measures for the Labelling of Artificial Intelligence-Generated and Synthetic Content	12
4.3.5.3 GB 45438-2025 Cybersecurity Technology - Labelling Method for Content Generated by Artificial Intelligence	12
4.3.6 EU: Artificial Intelligence Act.....	12
4.3.7 India NITI Aayog: Part I Principles for Responsible AI.....	12
4.3.8 Japan.....	13
4.3.8.1 Hiroshima AI Process: International Guiding Principles for Organizations Developing Advanced AI Systems	13
4.3.8.2 AI Guidelines for Business Version 1.0.....	13
4.3.9 South Korea - Framework Act on Artificial Intelligence Development and Establishment of a Foundation for Trustworthiness (AI Framework Act)	13
4.3.10 UK: AI Code of Practice.....	14
4.3.11 USA	14
4.3.11.1 California	14
4.3.11.1.1 AB-2013 Generative artificial intelligence: training data transparency	14
4.3.11.1.2 SB-942 California AI Transparency Act	14
4.3.11.1.3 AB-1836 and AB-2602.....	15
4.3.11.2 Colorado - Colorado Artificial Intelligence Act (CAIA)	15
4.3.11.3 Tennessee - Ensuring Likeness Image and Voice Security (ELVIS) Act	15
4.3.11.4 Utah - Artificial Intelligence Policy Act	16
5 Impact of GenAI on Intellectual Property Rights.....	16
5.1 Overview	16
5.2 Copyright theft and infringement	16
5.3 Understanding the Training Material	17
5.3.1 Data curation.....	17
5.3.1.1 Overview.....	17
5.3.1.2 Cleaning and filtering.....	17
5.3.1.3 Data annotation and labelling.....	18
5.4 Use of Open-Source Models	18
5.5 Purposeful Data Poisoning	18

6	Harmful Impacts from GenAI	19
6.1	Overview	19
6.2	Prompt Injection Attack	19
6.2.1	Overview	19
6.2.2	Direct Prompt Injection Attack	19
6.2.3	Indirect Prompt Injection Attack	19
6.3	Misinformation	20
6.4	GenAI Hallucinations	21
6.5	Loss of Confidentiality	22
6.6	Malicious Code Generation	22
6.7	Spam Generation	23
6.7.1	Overview	23
6.7.2	Phishing	23
6.7.3	Mitigations	23
6.8	Deepfakes	24
6.8.1	Overview	24
6.8.2	Detection and Prevention	25
6.8.3	Reporting and Removal	25
7	GenAI Content and Material	26
7.1	How GenAI is shared and spreads online	26
7.2	Best Practice Measures within GenAI Platforms / Services	26
7.2.1	Prevention by Design	26
7.2.2	Metadata	27
7.2.3	Red Teaming	28
7.3	Tackling the Content Shared from GenAI Platforms	28
7.3.1	Detection	28
7.3.2	Enforcement	29
7.3.4	Reporting	30
7.3.5	Removal	30
8	Conclusion	30
8.1	Overview	30
8.2	Trustworthy AI	31
8.2.1	Overview	31
8.2.2	Mapping of GenAI to Trustworthy AI	32
8.2.2.1	Overview	32
8.2.2.2	Table of Mapping	32
Annex A:	Change history	35
History		36

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: "*Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards*", which is available from the ETSI Secretariat. Latest updates are available on the [ETSI IPR online database](#).

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™**, **LTE™** and **5G™** logo are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM®** and the GSM logo are trademarks registered and owned by the GSM Association.

Foreword

This Technical Report (TR) has been produced by ETSI Technical Committee Securing Artificial Intelligence (SAI) (SRdAP).

ETSI TR 104 159 V1.1.1 (2026-01)

Modal verbs terminology

In the present document "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the [ETSI Drafting Rules](#) (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.