International Standard

ISO/IEC 25389

Information technology — The safe framework

First edition
2025-06

Reference number
ISO/IEC 25389:2025(en)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword ... **iv**
Introduction ... **v**
1 Scope ... **1**
2 Normative references ... **1**
3 Terms and definitions ... **1**
4 Digital trust and safety ... **3**
5 Commitments and practices ... **4**
5.1 Product development ... 5
5.2 Product governance ... 6
5.3 Product enforcement ... 6
5.4 Product improvement ... 7
5.5 Product transparency ... 7
6 Assessment framework ... **8**
6.1 Scoping ... 8
6.2 Tailoring ... 8
6.2.1 Tailoring methodology ... 8
6.2.2 Evaluating the organization's size and scale ... 9
6.2.3 Evaluating the impact of the product or digital service ... 10
6.2.4 Determine the initial recommended assessment level ... 10
6.2.5 Factor in additional business landscape considerations ... 11
6.3 Assessment Execution ... 12
6.3.1 Assessment Methodology ... 12
6.3.2 Discover ... 13
6.3.3 Identify ... 14
6.3.4 Assess ... 14
6.3.5 Test ... 15
6.3.6 Report ... 15
Annex A (informative) **Illustrative examples of the tailoring framework** ... **16**
Annex B (informative) **Risk Profile Questionnaire** ... **17**
Annex C (informative) **Summary of differences between L1, L2, and L3 Assessments** ... **19**
Annex D (informative) **Sample information discovery form** ... **20**
Annex E (informative) **Question Bank** ... **22**
Annex F (informative) **Illustrative example: product area report template** ... **25**
Bibliography ... **26**

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted (see www.iso.org/directives or www.iec.ch/members_experts/refdocs).

ISO and IEC draw attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO and IEC take no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO and IEC had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents and https://patents.iec.ch. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html. In the IEC, see www.iec.ch/understanding-standards.

This document was prepared by the Digital Trust & Safety Partnership (DTSP) (as The Safe Framework Specification) and drafted in accordance with its editorial rules. It was adopted, under the JTC 1 PAS procedure, by Joint Technical Committee ISO/IEC JTC 1, *Information technology*.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html and www.iec.ch/national-committees.

Introduction

Digital services are increasingly central to our daily lives, facilitating social discourse, economic activity, and much more. These services provide powerful tools for users across the globe to engage in a wide range of valuable online activity. But like any tool, they can also be misused to facilitate harmful behavior and content. Awareness of and action against this misuse has grown in recent years, which has led to increasing urgency in understanding, supporting, and evaluating effective ways to reduce harms associated with online content and behavior, while also protecting people's ability to express themselves, carry out business, access information, associate, work, study, and participate in their communities through digital services.

Striking this balance presents a considerable challenge. To begin, there is no one-size-fits-all approach to handling online content and associated behavioral risks or, more generally, to organizations' trust and safety operations. Depending on the nature of the digital service, each may face unique risks relative to the various products or features they provide – different threats, different vulnerabilities, and different consequences. Products or features may engage with end users directly or indirectly, as well as with other services or businesses. What is an effective practice for one digital service may not suit another, and highly prescriptive or rigid approaches to defining trust and safety practices are likely to be too broad, too narrow or have negative unintended consequences. Further, risks change over time and so approaches to mitigating them must also have room to evolve.

Given the diversity of digital services, it is important to define an overall framework and set of aims for what constitutes a responsible approach to managing content- and conduct-related risks, to which digital services can then map their specific practices. This flexible approach has been deployed in other domains, such as information security, yet existing frameworks are not sufficiently concrete to be applied when it comes to addressing harmful behavior and content online.

This document aims to fill this need by offering a framework of commitments to address content- and conduct-related risks. While the overarching commitments are uniform, the method by which they are fulfilled – whether by application of the illustrative practices in this document or alternatives – will vary by digital product or feature and evolve with both the challenges faced and advances made in the field of trust and safety.

This document also provides recommendations for organizations to evaluate the maturity of their implementation of these commitments through a rigorous and flexible approach to assessment.

This document is for the internal use of the organization responsible for trust and safety operations for a digital product or service. Recommendations for public reporting about the commitments and their implementation are outside the scope of this document.

This document is neither a management system standard, nor does it consider the issues of information security, privacy, and other aspects of security, privacy, and data management that are addressed by existing international standards.