

BS 8611:2023



BSI Standards Publication

Robots and robotic devices – Ethical design and application of robots and robotic systems – Guide



Publishing and copyright information

The BSI copyright notice displayed in this document indicates when the document was last issued.

© The British Standards Institution 2023

Published by BSI Standards Limited 2023

ISBN 978 0 539 21195 5

ICS 25.040.30

The following BSI references relate to the work on this document:

Committee reference AMT/10/1

Draft for comment 22/30451737 DC

Amendments/corrigenda issued since publication

Date	Text affected
------	---------------

Contents

	Page
Foreword	II
1 Scope	1
2 Normative references	1
3 Terms and definitions	1
4 Ethical risks	2
4.1 General	2
<i>Table 1 — Ethical issues, hazards and risks</i>	3
4.2 Ethical hazard identification	8
4.3 Ethical risk assessment	8
4.4 Learning robots	9
<i>Figure 1 — Ethical risk assessment and management process</i>	10
5 Ethical guidelines and measures	11
5.1 General societal ethical guidelines	11
5.2 Application of ethical recommendations and measures	13
<i>Table 2 — Ethical design parameters</i>	14
5.3 Commercial/financial ethical guidelines and measures	14
5.4 Environmental ethical guidelines, measures and sustainable development	15
6 Ethics-related system design recommendations	15
6.1 General principles for ethical design	15
6.2 Protection against the perception of harm	16
7 Verification and validation	16
7.1 General	16
7.2 Mitigation	17
7.3 Suggested methods	17
8 Information for use	19
8.1 General	19
8.2 Markings or indications	20
8.3 User manual	21
8.4 Service manual	21
Bibliography	22

Summary of pages

This document comprises a front cover, an inside front cover, pages I to IV, pages 1 to 23, an inside back cover and a back cover.

Foreword

Publishing information

This British Standard is published by BSI Standards Limited, under licence from The British Standards Institution, and came into effect on 31 March 2023. It was prepared by Subcommittee AMT/10/1, *Ethics for robots and autonomous systems*, under the authority of Technical Committee AMT/10, *Robotics*. A list of organizations represented on these committees can be obtained on request to the committee manager.

Supersession

This British Standard supersedes BS 8611:2016, which is withdrawn.

Information about this document

This is a full revision of the standard, and introduces the following principal changes:

- the addition of new ethical hazards in [Table 1](#); and
- additional guidance on the application of the standard including a new flowchart for ethical risk assessment and management ([Figure 1](#)) and new guidance on ethical design parameters ([Table 2](#)).

This British Standard addresses issues in a field of technology that is rapidly changing and influencing society. Although efforts have been made to incorporate the anticipated effects of future developments in the standard, it is not possible for all future circumstances to have been considered or covered, for example non-embodied autonomous systems.

This publication can be withdrawn, revised, partially superseded or superseded. Information regarding the status of this publication can be found in the Standards Catalogue on the BSI website at bsigroup.com/standards, or by contacting the Customer Services team.

Where websites and webpages have been cited, they are provided for ease of reference and are correct at the time of publication. The location of a webpage or website, or its contents, cannot be guaranteed.

Use of this document

As a guide, this British Standard takes the form of guidance and advisory recommendations. It is not to be quoted as if it were a specification or a code of practice.

It has been assumed in the preparation of this British Standard that the execution of its provisions will be entrusted to appropriately qualified and experienced people, for whose use it has been produced.

Presentational conventions

The guidance in this document is presented in roman (i.e. upright) type. Any recommendations are expressed in sentences in which the principal auxiliary verb is “should”.

Additional commentary, explanation and general informative material is presented in smaller italic type.

Where words have alternative spellings, the preferred spelling of the Shorter Oxford English Dictionary is used (e.g. “organization” rather than “organisation”).

Contractual and legal considerations

This publication has been prepared in good faith, however no representation, warranty, assurance or undertaking (express or implied) is or will be made, and no responsibility or liability is or will be accepted by BSI in relation to the adequacy, accuracy, completeness or reasonableness of this publication. All and any such responsibility and liability is expressly disclaimed to the full extent permitted by the law.

This publication is provided as is, and is to be used at the recipient's own risk.

The recipient is advised to consider seeking professional guidance with respect to its use of this publication.

This publication is not intended to constitute a contract. Users are responsible for its correct application.

Compliance with a British Standard cannot confer immunity from legal obligations.

1 Scope

This British Standard gives guidance on the identification of potential ethical harm and provides guidelines on the design control of robots through their whole lifecycle. It builds on existing safety requirements for different types of robots (e.g. industrial, personal care, service and medical).

This British Standard describes ethical hazards associated with robots and provides guidance to eliminate or reduce the risks associated with these ethical hazards. Significant ethical hazards are presented and guidance is given on how these hazards can be addressed for various robots, where possible.

Ethical hazards are broader than physical hazards; they additionally include psychological and social hazards (physical hazards imply ethical hazards and safety design features are part of ethical design). Safety elements are covered by safety standards; this British Standard is concerned with ethical elements.

This British Standard is primarily intended for use by designers and managers.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes provisions, or limits the application, of this document.¹⁾ For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

BS ISO 8373, *Robotics – Vocabulary*

3 Terms and definitions

For the purposes of this British Standard, the terms and definitions given in BS ISO 8373 and the following apply.

3.1 autonomous system

system which has the ability to perform intended tasks based on current state, knowledge and sensing, without human intervention

3.2 ethical behaviour

observance of/adherence to respect and care for the community of life, ecological integrity, social and cultural values, economic equity, non-violence and peace

3.3 ethical harm

anything likely to compromise psychological and/or societal and environmental well-being

NOTE Examples of ethical harm include stress, embarrassment, anxiety, addiction, discomfort, deception that induces negative consequences, humiliation, being disregarded, social isolation.

3.4 ethical hazard

potential source of ethical harm

3.5 ethical risk

probability of ethical harm occurring from the frequency and severity of exposure to a hazard

¹⁾ Documents that are referred to solely in an informative manner are listed in the Bibliography.

3.6 ethics

common understanding of principles that constrain and guide human behaviour

3.7 harm

injury or damage to health or well-being

3.8 system

set of parts which, when interoperating, has behaviour that is not present in any of the parts themselves

NOTE 1 This definition includes systems of fleeting duration, such as cars passing each other and behaving appropriately during this moment, to long-lived systems such as Voyager II.

NOTE 2 In a networked world, the robot is unlikely to be instantiated as a stand-alone device. It is more likely that it would be a component of a system, or a larger system of systems, in which the robot becomes a physical prosthesis in executing the larger system's purposes, with supervisory control of the robot exercised elsewhere, perhaps in some foreign jurisdiction. This British Standard does not consider the ethical complexities introduced by this scenario; its focus is on the ethics of interactions over time between a robot and a human being, bearing in mind that there are significant differences between individual humans in needs, capabilities, behaviours and understanding.

3.9 vulnerable people

people having reduced physical, sensory, societal, economical or mental capabilities (e.g. partially disabled, older adults having some reduction in their physical or cognitive capabilities due to age-related impairments or long-term conditions), or lack of experience and knowledge (e.g. older children)

[SOURCE: BS EN 60335-1:2012+A15:2021, 3.Z.5, modified]

4 Ethical risks

4.1 General

Ethical issues can be divided into societal, application, commercial/financial and environmental, as described in [Clause 5](#) to [Clause 8](#). [Table 1](#) gives examples of the ethical hazards and risks associated with these issues. BS ISO 31000 gives further guidance.

Table 1 — *Ethical issues, hazards and risks*

Ethical issue	Ethical hazard	Ethical risk	Mitigation	Comment	Verification/ validation
Societal	Loss of trust (human robot)	Robot no longer used or is misused, abused	Design to enable reliability in behaviour	If unexpected behaviour occurs, encourage traceability to help explain what happened	User validation
	Deception (intentional or unintentional)	Confusion, unintended interactions and consequences (perhaps delayed), eventual loss of trust	Avoid deception due to the behaviour and/or appearance of the robot and enable transparency and traceability of robotic nature and behaviour	—	Software verification; user validation; expert guidance
	Anthropomorphization	Misinterpretation	Avoid unjustifiable anthropomorphization	See Deception (above)	User validation; expert guidance
			Clarification of intent to simulate human or not, or intended or expected behaviour	Use anthropomorphization only for well-defined, limited and socially-accepted purposes	
Lack of privacy and confidentiality	Unauthorized access, collection and/or distribution of data, e.g. coming into the public domain or to unauthorized, unwarranted entities	Clarity of function	Control of data, justification of data collection and distribution	Privacy by design	Software verification
			Enable user awareness of data management and obtain informed consent in appropriate contexts, in compliance with local data protection regulations	Data encryption, storage location, adherence to legislation	
Lack of respect for cultural diversity and pluralism	Loss of trust in the device, embarrassment, shame, offence	Awareness of cultural norms incorporated into programming	Organizational, professional, regional		Software verification; user validation