International Standard

ISO 26324

Information and documentation — Digital object identifier system

Information et documentation — Système d'identifiant numérique d'objet

Third edition
2025-03

Reference number
ISO 26324:2025(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

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

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 ISO document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

ISO draws attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO takes 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 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. ISO 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.

This document was prepared by Technical Committee ISO/TC 46, *Information and documentation*, Subcommittee SC 9, *Identification and description*.

This third edition cancels and replaces the second edition (ISO 26324:2022), which has been technically revised.

The main changes are as follows:

— case-sensitiveness of DOI names is clarified;

— the list of common representations of DOI names is updated;

— the specification of, and procedures associated with, the system metadata associated with each DOI Name (see Annex B) are revised to reflect current practice;

— Annex D is replaced with a summary of changes across editions of this document.

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.

Introduction

The digital object identifier (DOI®[1]) system, which was first deployed in 1998, provides an infrastructure for persistent unique identification of objects of any type.

DOI is an initialism for “digital object identifier”, meaning a “digital identifier of an object” rather than an “identifier of a digital object”: DOI is not limited to digital objects and can be used to identify any abstract, physical or non-physical object. In this document, the term “digital object identifier” refers to the system defined in this document, unless otherwise stated.

Objects within the DOI system, i.e. referents, are identified by their DOI names. A DOI name is an opaque string, which does not have a discernible meaning on its own and is for use by humans and machines alike. Each DOI name is registered by a registrant.

To support diverse applications worldwide, a DOI name uses characters from the Unicode character set. More specifically, it can use any Unicode character intended to be written, printed, or otherwise displayed in a form that can be read by humans – formally referred to as a Graphic character.

With this flexibility comes ambiguities when representing or exchanging DOI names. For example:

— the character "Á" (Latin Capital Letter A with Acute) can be encoded either on its own or as the character "A" (Latin Capital Letter A) followed by the combining character ◌́ (Combining Acute Accent);

— multiple schemes (UTF-8, UTF-16 or UTF-32) can be used when serializing a DOI name to bytes for interchange between machines;

— the glyph " Å " can either correspond to the Angstrom Sign or the Latin Capital Letter A with Ring Above.

To avoid these pitfalls, this document specifies the syntax of a DOI name as a sequence of Unicode code points, where each code point is an integer between 0 and 0x10FFFF, and the fundamental unit of encoding in Unicode; and describes several representations appropriate for interchange and visual representation.

The assignment and administration of DOI names is handled by the DOI system, which offers a useful set of functionalities, including:

— persistence, if material is moved, rearranged, or bookmarked,

— interoperability with other data from other sources,

— extensibility by adding new features and services through management of groups of DOI names,

— single management of data for multiple output formats (platform independence),

— class management of applications and services, and

— dynamic updating of metadata, applications and services.

The DOI system is designed to work over the Internet. A DOI name is permanently assigned to an object to provide a resolvable persistent network link to current information about that object, including where the object, or information about it, can be found on the Internet. While information about an object can change over time, its DOI name will not change. A DOI name can be resolved within the DOI system to information related to the identified object such as the location of metadata or a location of the object.

The DOI system enables the construction of automated services and transactions. Applications of the DOI system include, but are not limited to, managing information and documentation location and access;

1) DOI® is a registered trademark. The DOI Handbook published by the ISO 26324 Registration Authority (see Clause 8) contains information on trademark issues. The name and contact information of the Registration Authority for this document can be found at https://www.iso.org/maintenance_agencies.html.

managing metadata; facilitating electronic transactions; persistent unique identification of any form of any data; and commercial and non-commercial services.

An object associated with a DOI name is described unambiguously by system metadata, to support identification and description. The data model supports interoperability between applications.

The scope of the DOI system is not defined by reference to the type of content (format, etc.) of the referent, but by reference to the functionalities it provides and the context of use. The DOI system provides, within networks of DOI applications, for unique identification, persistence, resolution, metadata and interoperability with other identifier schemes.

Information and documentation — Digital object identifier system

1 Scope

This document specifies the syntax, description and resolution functional components of the digital object identifier system. It specifies the general principles for the creation, registration and administration of DOI names (where DOI is an initialism for “digital object identifier”).

This document defines the syntax for a DOI name, which is used for the identification of an object of any material form (non-physical or physical) or an abstraction (such as a textual work) where there is a functional need to distinguish it from other objects.

The DOI name does not replace, nor is it an alternative for, an identifier used in another scheme, such as the schemes defined by ISO/TC 46/SC 9. This document describes how the DOI system can be used in conjunction with another identifier scheme (for example, to provide additional functionality, such as resolution, where this is not already available), and how the character string of that other scheme can be integrated into the DOI system through system metadata or the DOI syntax or both.

This document does not specify particular technologies to implement the syntax, description and resolution functional components of the digital object identifier system.

2 Normative references

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

ISO/IEC 10646, *Information technology — Universal coded character set (UCS)*

Namespace Registration for Digital Object Identifier (DOI), IANA Registry of URN Namespaces, https://www.iana.org/assignments/urn-formal/doi

3 Terms and definitions

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

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

3.1
DOI system
social and technical infrastructure for the assignment and administration of *DOI names* (3.3) as identifiers in computer-readable form through assignment, resolution, referent description, administration, etc.

3.2
code point
any value in the Unicode codespace

Note 1 to entry: In running text, an individual Unicode code point is expressed as U+n, where n is four to six hexadecimal digits, using the digits 0–9 and uppercase letters A–F (for 10 through 15, respectively); and a formal Unicode name is shown in small capitals.

[SOURCE: ISO/IEC 10646:2020, 3.9]

3.3
DOI name
opaque string (3.9) that specifies a unique object within the *DOI system* (3.1)

Note 1 to entry: A DOI name consists of a sequence of code points specified by the *DOI syntax* (3.5).

Note 2 to entry: The terms "identifier" and "number" are sometimes but not always used in the same sense and are to be avoided where ambiguity can arise. The unqualified use of "DOI" alone can also be ambiguous. Therefore "DOI" is always used in conjunction with a specific noun such as DOI name or *DOI system* (3.1) unless the meaning is sufficiently clear from an earlier mention or the specific context.

3.4
object
entity that can become a referent in the scope of the *DOI system* (3.1), including, but not limited to, digital, non-physical, physical, and abstract entities

Note 1 to entry: Digital, physical or abstract forms of an entity can be of relevance in information and documentation (e.g. resources, people or agreements).

Note 2 to entry: A particular object identified by a specific DOI name is the referent of that *DOI name* (3.3).

Note 3 to entry: Examples of objects include, but are not limited to: a novel as an abstract work, a specific version of a television episode, a specific scene within a version of a movie, a ceremonial robe in a museum collection, a product offering such as building insulation, a batch of manufactured products, an instance of a product, a person, a source file, a binary software package, a scientific journal, a research grant, a project, etc.

3.5
DOI syntax
rules for the form and sequence of *code points* (3.2) comprising any *DOI name* (3.3), specifically the form and sequence of *code points* (3.2) of a prefix element, separator and suffix element

3.6
directory indicator
unique *opaque string* (3.9) allocated to a *registrant* (3.10) for the purpose of assignment of *DOI names* (3.3), forming part of the prefix element of the *DOI syntax* (3.5) but having no other implied meaning

3.7
registrant code
unique *opaque string* (3.9) forming part of the prefix element of the *DOI syntax* (3.5) but having no other implied meaning

3.8
system metadata
specific data associated with the referent of a *DOI name* (3.3), based on a data model that enables the *referent* (3.15) to be associated with data of any desired degree of precision and granularity to support identification and description

Note 1 to entry: system metadata is specified in Annex B.

3.9
opaque string
sequence of *code points* (3.2) that has no meaning discernible by simple inspection

Note 1 to entry: To discover meaning, there is a need to refer to metadata.

3.10
registrant
person or organization that has requested and received the registration of a particular *DOI name* (3.3)

3.11
interoperability
ability of independent systems to exchange meaningful information and initiate actions from each other, in order to operate together to mutual benefit

Note 1 to entry: In particular, interoperability constitutes the ability for loosely-coupled independent systems to be able to collaborate and communicate. See Paskin (2006) for further information about interoperability.

3.12
resolution
process of submitting a *DOI name* (3.3) to a service and receiving in return a *resolution record* (3.13)

Note 1 to entry: This can involve one or more intermediate mapping operations. The resolution might or might not return an instance of the *referent* (3.15). Multiple resolution is the simultaneous return as output of several pieces of current information related to the referent.

3.13
resolution record
data that is the response to a *resolution* (3.12) request providing information related to the *referent* (3.15)

Note 1 to entry: A resolution record does not necessarily include the referent or associated system metadata, but instead can include the location of such information.

3.14
resolution service
capability provided by the *DOI system* (3.1) that performs *resolution* (3.12)

3.15
referent
particular object identified by a *DOI name* (3.3)

3.16
unique identification
specification by a *DOI name* (3.3) of one and only one *referent* (3.15)

3.17
persistent
existence, and ability to be used in services outside the direct control of the issuing assigner, without a stated time limit

3.18
first class
having an identity of itself, not as some attribute of an object

Note 1 to entry: An address is an attribute of something, whereas the thing that has this attribute is a first class object. A *DOI name* (3.3) references an entity as a first-class object, not simply the place where the object is located. It may then resolve to a location.

4 DOI name

4.1 Syntax

4.1.1 General characteristics

A DOI name shall consist of an ordered sequence of code points of the Graphic type, as specified in ISO/IEC 10646.

NOTE 1 The Unicode Graphic type includes all code points that are letter, mark, number, punctuation, symbol and spaces. It excludes, for example, control code points such as U+0009 HORIZONTAL TABULATION.

NOTE 2 The term code point is used instead of the term character, which is ambiguous in the context of Unicode where a given abstract character can be encoded in multiple ways (see Example 3).

The code points are arranged in a DOI prefix and a DOI suffix separated by U+002F SOLIDUS.

NOTE 3 U+002F SOLIDUS is also referred to as forward slash ("/").

There is no defined limit on the length of the DOI name, or of the DOI prefix or DOI suffix.

Further constraints on code points (e.g. use of language-specific alphanumeric characters) can be defined for an application by the ISO 26324 Registration Authority.

The combination of a unique DOI prefix (assigned to a particular DOI registrant) and a unique DOI suffix (provided by that registrant for a specific object) shall be unique. This allows the de-centralized allocation of DOI names. The registration of the combination of the prefix and suffix in the DOI system also serves to validate the DOI syntax for a given DOI name.

The DOI name shall be regarded as an opaque string by users of the DOI system. No definitive information shall be inferred from the specific sequence of code points that make up a DOI name. In particular, the inclusion in a DOI name of any DOI prefix allocated to a specific registrant does not provide evidence of the ownership of rights or current management responsibility of any intellectual property in the referent. The mere assignment of a DOI name to an object shall not imply endorsement of the object by any party.

When comparing two DOI names for equivalence, no normalization, as defined in ISO/IEC 10646, shall be performed and the DOI names are equivalent if, and only if, their code point sequences are identical, except that a code point in the range U+0041..U+005A (corresponding to LATIN CAPITAL LETTER A to LATIN CAPITAL LETTER Z) is considered identical to the corresponding code point in the range U+0061..U+007A (corresponding to characters LATIN SMALL LETTER A to LATIN SMALL LETTER Z).

NOTE 4 The rule above has the effect of making DOI names case-insensitive only when testing for equivalence and only with respect to the Basic Latin Unicode block. It does not restrict DOI names to containing only uppercase or lowercase letters.

EXAMPLE 1 The following DOI names are equivalent because U+0053 LATIN CAPITAL LETTER S and U+0073 LATIN SMALL LETTER S are considered identical:

10.5594/SMPTE.ST2067-21.2020

10.5594/sMPTE.sT2067-21.2020

EXAMPLE 2 The following DOI names are not equivalent because U+00C1 LATIN CAPITAL LETTER A WITH ACUTE and U+00E1 LATIN SMALL LETTER A WITH ACUTE are not considered identical:

10.26321/Á.GUTIÉRREZ.ZARZA.02.2018.03

10.26321/á.gutiérrez.zarza.02.2018.03

EXAMPLE 3 The code point sequences <U+00C1> and <U+0041, U+0301> are not identical even though they are both rendered identically as the abstract character "Á". As a result, the following DOI names, expressed as a sequence of Unicode code points, are not equivalent, even though they are both rendered as "10.26321/Á.GUTIÉRREZ.ZARZA.02.2018.03":