ISO/~~DIS~~PRF 24138:~~2024(E)~~

ISO/TC 46/SC 9

Secretariat: ANSI

Date: 2024-~~02-05~~03-13

Information and documentation — International Standard Content Code (ISCC)

iTeh Standards
(https://standards.iteh.ai)
Document Preview

PROOF

ISO/PRF 24138
https://standards.iteh.ai/catalog/standards/iso/4bd4b8ba-b648-409e-8713-22cb3cbbe602/iso-prf-24138

Edited DIS - MUST BE USED FOR FINAL DRAFT

Style Definition ...
Formatted: Font: Bold, English (United Kingdom)
Formatted: Right: 1.5 cm, Bottom: 1 cm, Gutter: 0 cm, Header distance from edge: 1.27 cm, Footer distance from edge: 1.27 cm
Formatted ...
Formatted: zzCover large
Formatted: English (United Kingdom)
Formatted: English (United Kingdom)
Formatted: Font: Bold, English (United Kingdom)
Formatted: Cover Title_A1
Formatted: English (United Kingdom)

© ISO 2024

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

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

Published in Switzerland

iTeh Standards
(https://standards.iteh.ai)
Document Preview

ISO/PRF 24138
https://standards.iteh.ai/catalog/standards/iso/4bd4b8ba-b648-409e-8713-22cb3cbbe602/iso-prf-24138

Formatted: Font: 11 pt, Bold
Formatted: Font: 11 pt, Bold
Formatted: Font: Bold
Formatted: HeaderCentered
Formatted: Font: 11 pt, Bold
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: French (Switzerland)
Formatted: French (Switzerland)
Formatted: French (Switzerland)
Formatted: English (United Kingdom)
Formatted: English (United Kingdom)
Formatted: FooterPageRomanNumber

Contents

Foreword ix
Introduction x
1 Scope 1
2 Normative references 1
3 Terms and definitions 1
4 Structure and format of the ISCC 4
4.1 General structure 4
4.2 ISCC-HEADER 7
4.2.1 General 7
4.2.2 MainTypes 8
4.2.3 SubTypes 9
4.2.4 Version 9
4.2.5 Length 9
4.3 ISCC-BODY 11
4.4 Encoding 11
4.4.1 Canonical form 11
4.4.2 URI encoding 11
4.4.3 Multiformats encoding 12
4.4.4 Readable encoding 12
5 ISCC-UNITs 13
5.1 Meta-Code 13
5.1.1 General 13
5.1.2 Purpose 13
5.1.3 Format 13
5.1.4 Inputs 14
5.1.5 Outputs 16
5.1.6 Seed metadata processing 16
5.1.7 Metadata embedding 18
5.1.8 Metadata extraction 19
5.2 Content-Codes 19
5.2.1 General 19
5.2.2 Purpose 20
5.3 Content-Code Subtype Text 20
5.3.1 General 20
5.3.2 Format 20

iTeh Standards
(https://standards.iteh.ai)
Document Preview
ISO/PRF 24138
https://standards.iteh.ai/catalog/standards/iso/4bd4b8ba-b648-409a-8717-22cb3cbbe602/iso-prf-24138

~~© ISO 2024 – All rights reserved~~

© ISO 2024 – All rights reserved

Edited DIS - MUST BE USED FOR FINAL DRAFT

Formatted: Font: 11 pt, Bold
Formatted: HeaderCentered, Left
Formatted: Font: 11 pt, Bold
Formatted: Font: 11 pt, Bold
Formatted: Font: Bold
Formatted: Tab stops: Not at 17.2 cm
Formatted: FooterPageRomanNumber

5.3.3 Inputs 21
5.3.4 Outputs 21
5.3.5 Processing 22
5.3.6 Conformance 22
5.4 Content-Code Subtype Image 22
5.4.1 General 22
5.4.2 Format 23
5.4.3 Inputs 23
5.4.4 Outputs 23
5.4.5 Processing 24
5.4.6 Conformance 24
5.5 Content-Code Subtype Audio 25
5.5.1 General 25
5.5.2 Format 25
5.5.3 Inputs 25
5.5.4 Outputs 26
5.5.5 Processing 26
5.5.6 Conformance 26
5.6 Content-Code Subtype Video 26
5.6.1 General 26
5.6.2 Format 27
5.6.3 Inputs 27
5.6.4 Outputs 27
5.6.5 Processing 28
5.6.6 Conformance 28
5.7 Content-Code Subtype Mixed 28
5.7.1 General 28
5.7.2 Format 29
5.7.3 Inputs 29
5.7.4 Outputs 29
5.7.5 Processing 30
5.7.6 Conformance 30
5.8 Data-Code 30
5.8.1 General 30
5.8.2 Format 31
5.8.3 Inputs 31
5.8.4 Outputs 31

iTeh Standards
(https://standards.iteh.ai)
Document Preview
ISO/PRF 24138
https://standards.iteh.ai/catalog/standards/iso/4bd4b8ba-b648-409e-87...22cb3cbbe602/iso-prf-24138

© ISO 2024 – All rights reserved

Formatted: Font: 11 pt, Bold
Formatted: Font: 11 pt, Bold
Formatted: HeaderCentered
Formatted: Font: 11 pt, Bold
Formatted: Font: Bold
Formatted: FooterPageRomanNumber

5.8.5 Processing 31
5.8.6 Conformance 32
5.9 Instance-Code 32
5.9.1 General 32
5.9.2 Format 34
5.9.3 Inputs 35
5.9.4 Outputs 35
5.9.5 Processing 35
5.9.6 Conformance 35
6 ISCC-CODE 36
6.1 General 36
6.2 Purpose 36
6.3 Format 36
6.3.1 General 36
6.3.2 SubTypes for ISCC-CODEs 36
6.3.3 Length and composition of ISCC-CODEs 37
6.4 Inputs 37
6.5 Outputs 38
6.6 Processing 38
6.7 Comparing ISCC-CODEs 39
6.8 Conformance 39
Annex A (normative) Relationship between ISCC and other identifier systems 40
Annex B (normative) ISCC metadata 43
Annex C (informative) Evolution of this document 46
Annex D (normative) Reference implementation 47
Bibliography 48

Foreword vi
Introduction vii
1 Scope 1
2 Normative references 1
3 Terms and definitions 2
4 Structure and format of the ISCC 5
4.1 General structure 5
4.2 ISCC-HEADER 6
4.2.1 MainTypes 7

iTeh Standards
(https://standards.iteh.ai)
Document Preview
ISO/PRF 24138
https://standards.iteh.ai/catalog/standards/iso/4bd4b8ba-b648-409e-87f1-22cb3cbbe602/iso-prf-24138

© ISO 2024 – All rights reserved

Edited DIS - MUST BE USED FOR FINAL DRAFT

Formatted: Font: 11 pt, Bold
Formatted: HeaderCentered, Left
Formatted: Font: 11 pt, Bold
Formatted: Font: 11 pt, Bold
Formatted: Font: Bold
Formatted: FooterPageRomanNumber

4.2.2 SubTypes 7
4.2.3 Version 8
4.2.4 Length 8
4.3 ISCC-BODY 9
4.4 Encoding 9
4.4.1 Canonical form 9
4.4.2 ISCC:KEC43HJLPUSHVAZT66YLPUWNVACWYPIV533TRQMWF2IUQYSP5LA4CTYURI encoding 9
4.4.3 Multiformats encoding 9
4.4.4 Readable encoding 10
5 ISCC-UNITs 11
5.1 Meta-Code 11
5.1.1 General 11
5.1.2 Purpose 11
5.1.3 Format 11
5.1.4 Inputs 11
5.1.5 Outputs 13
5.1.6 Seed metadata processing 13
5.1.7 Metadata embedding 15
5.1.8 Metadata extraction 15
5.2 Content-Codes 16
5.2.1 General 16
5.2.2 Purpose 16
5.3 Content-Code Subtype Text 16
5.3.1 General 16
5.3.2 Format 16
5.3.3 Inputs 17
5.3.4 Outputs 17
5.3.5 Processing 17
5.3.6 Conformance 17
5.4 Content-Code Subtype Image 18
5.4.1 General 18
5.4.2 Format 18
5.4.3 Inputs 18
5.4.4 Outputs 18
5.4.5 Processing 19
5.4.6 Conformance 19
5.5 Content-Code Subtype Audio 20

iTeh Standards
(https://standards.iteh.ai)
Document Preview
ISO/PRF 24138
https://standards.iteh.ai/catalog/standards/iso/4bd4b8ba-b648-409e-8713-22cb3cbbe602/iso-prf-24138

© ISO 2024 – All rights reserved

Formatted: Font: 11 pt, Bold
Formatted: Font: 11 pt, Bold
Formatted: HeaderCentered
Formatted: Font: 11 pt, Bold
Formatted: Font: Bold
Formatted: FooterPageRomanNumber

5.5.1 General 20
5.5.2 Format 20
5.5.3 Inputs 20
5.5.4 Outputs 20
5.5.5 Processing 20
5.5.6 Conformance 21
5.6 Content-Code Subtype Video 21
5.6.1 General 21
5.6.2 Format 21
5.6.3 Inputs 21
5.6.4 Outputs 21
5.6.5 Processing 22
5.6.6 Conformance 22
5.7 Content-Code Subtype Mixed 22
5.7.1 General 22
5.7.2 Format 23
5.7.3 Inputs 23
5.7.4 Outputs 23
5.7.5 Processing 23
5.7.6 Conformance 23
5.8 Data-Code 24
5.8.1 General 24
5.8.2 Format 24
5.8.3 Inputs 24
5.8.4 Outputs 24
5.8.5 Processing 24
5.8.6 Conformance 25
5.9 Instance-Code 25
5.9.1 General 25
5.9.2 Format 26
5.9.3 Inputs 27
5.9.4 Outputs 27
5.9.5 Processing 27
5.9.6 Conformance 27
6 ISCC-CODE 27
6.1 General 27
6.2 Purpose 27

iTeh Standards
(https://standards.iteh.ai)
Document Preview
ISO/PRF 24138
https://standards.iteh.ai/catalog/standards/iso/4bd4b8ba-b648-409e-8713-22cb3cbbe602/iso-prf-24138

© ISO 2024 – All rights reserved

Edited DIS - MUST BE USED FOR FINAL DRAFT

Formatted: Font: 11 pt, Bold
Formatted: HeaderCentered, Left
Formatted: Font: 11 pt, Bold
Formatted: Font: 11 pt, Bold
Formatted: Font: Bold
Formatted: FooterPageRomanNumber

6.3 Format 27
6.3.1 SubTypes for ISCC-CODEs 28
6.3.2 Length and composition of ISCC-CODEs 28
6.4 Inputs 28
6.5 Outputs 29
6.6 Processing 29
6.7 Comparing ISCC-CODEs 29
6.8 Conformance 29
Annex A (normative) Relationship between ISCC and other identifier systems 30
A.1 Principles 30
A.2 Linking ISCCs to other identifiers 30
A.3 Use of ISCC by other identifier schemes 30
Annex B (normative) ISCC metadata 32
B.1 General 32
B.2 Seed metadata 32
B.3 Technical metadata 33
B.4 Other metadata 33
Annex C (informative) Evolution of this document 34
Annex D (normative) Reference implementation 35
Bibliography 36

iTeh Standards
(https://standards.iteh.ai)
Document Preview
ISO/PRF 24138
https://standards.iteh.ai/catalog/standards/iso/4bd4b8ba-b648-409e-8713-22cb3cbbe602/iso-prf-24138

© ISO 2024 – All rights reserved

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.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.*

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.htmlwww.iso.org/members.html.

iTeh Standards (https://standards.iteh.ai) Document Preview
ISO/PRF 24138

Formatted: Font: 11 pt, Bold
Formatted: Font: 11 pt, Bold
Formatted: Font: 11 pt, Bold
Formatted: Font: Bold
Formatted: HeaderCentered, Left
Formatted: Adjust space between Latin and Asian text, Adjust space between Asian text and numbers
Formatted: English (United Kingdom)
Formatted: English (United Kingdom)
Field Code Changed
Formatted: Adjust space between Latin and Asian text, Adjust space between Asian text and numbers
Formatted: English (United Kingdom)
Formatted: English (United Kingdom)
Formatted: FooterPageRomanNumber

© ISO 2024 – All rights reserved

Edited DIS - MUST BE USED FOR FINAL DRAFT

Introduction

While ISO/TC 46/SC 9 has established a variety of specific identifier standards, a content-dependent identifier for digital assets in all content formats has not yet been agreed.

Digital content is dynamic, always in motion, and acted upon globally by a variety of entities with different interests and requirements. Digital content continuously re-encodes, resizes, and re-compresses, changing its data as it travels through a complex network of actors and systems.

The International Standard Content Code (ISCC) is an identifier for numerous types of digital assets. An ISCC-CODE is generated from the digital content itself. It is the result of processing the digital content using a variety of algorithms including hash algorithms. The generated ISCC-CODE supports data integrity verification and preserves an estimate of the data, digital content and metadata similarity. However, ISCC has different functionality from content recognition systems.

The ISCC supports the association of higher-level identifiers (like work and product identifiers) with the digitally encoded manifestations of content. The ISCC does not specify a system for managing authoritative metadata. Other content identifier standards can use ISCC to support discoverability of their identifiers and metadata based on digital content.

Organizations, individuals and machines may generate ISCCs for numerous kinds of digital assets and use them for identification and management of those assets.

ISCCs are neither manually nor automatically assigned to digital media assets. Instead, ISCCs are derived from media assets according to the procedures described in this document. Unrelated parties can independently derive the same ISCC from a given media asset.

ISCCs exclusively reference media assets without any implication about ownership. As such, ISCCs are not managed authoritatively by any institution or entity.

The ISCC enables interoperability between different actors and systems using digital assets and supports scenarios that require content deduplication, database synchronization and indexing, integrity verification, timestamping, versioning, data provenance, similarity clustering, anomaly detection, usage tracking, allocation of royalties, fact-checking and general digital asset management use-cases.

This document includes sections targeting a general audience but also descriptions of more technical procedures.

Future editions of this document can be developed as outlined in ~~Annex C.~~Annex C.

iTeh Standards
(https://standards.iteh.ai)
Document Preview
ISO/PRF 24138
https://standards.iteh.ai/catalog/standards/iso/...b648-409e-8713-22cb3cbbe602/iso-prf-24138

~~© ISO 2024 – All rights reserved~~

© ISO 2024 – All rights reserved

Formatted: Font: 11 pt, Bold
Formatted: Font: 11 pt, Bold
Formatted: HeaderCentered
Formatted: Font: 11 pt, Bold
Formatted: Font: Bold
Formatted: Adjust space between Latin and Asian text, Adjust space between Asian text and numbers
Formatted: FooterPageRomanNumber

Information and documentation — International Standard Content Code (ISCC)

1 Scope

This document specifies the syntax and structure of the International Standard Content Code (ISCC), as an identification system for digital assets (including encodings of text, images, audio, video or other content across all media sectors). It also describes ISCC metadata and the use of ISCC in conjunction with other schemes, such as DOI, ISAN, ISBN, ISRC, ISSN, and ISWC.

An ISCC applies to a specific digital asset and is a data-descriptor deterministically constructed from multiple hash digests using the algorithms and rules in this document. This document does not provide information on registration of ISCCs.

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 10646:2020, Information technology — Universal coded character set (UCS)

ISO/IEC 15938, Information technology — Multimedia content description interface

ISO/IEC 21778:2017, Information technology — The JSON data interchange syntax

~~ISO/IEC 10646:2020, Information technology — Universal coded character set (UCS)~~

IETF RFC 4648, *The Base16, Base32, and Base64 Data Encodings*[1]

IETF RFC 2397, *The "data" URL scheme*[2]

IETF RFC 8785, *JSON Canonicalization Scheme (JCS)*[3]

W3C, C14N 1.1, *Canonical XML Version 1.1*[4]

W3C, JSON-LD 1.1, *A JSON-based Serialization for Linked Data*[5]

3 Terms and definitions

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

[1] Online available: https://datatracker.ietf.org/doc/html/rfc4648
[2] Online available: https://datatracker.ietf.org/doc/html/rfc2397
[3] Online available: https://datatracker.ietf.org/doc/html/rfc8785
[4] ~~Online available: https://www.w3.org/TR/xml-c14n11~~ Online available: https://www.w3.org/TR/xml-c14n11
[5] Online available: https://www.w3.org/TR/json-ld/

© ISO 2024 – All rights reserved

iTeh Standards (https://standards.iteh.ai) Document Preview ISO/PRF 24138 https://standards.iteh.ai/catalog/standards/iso/4bd4b8ba-b648-409e-8713-2...

Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Adjust space between Latin and Asian text, Adjust space between Asian text and numbers, Tab stops: Not at 0.7 cm + 1.4 cm + 2.1 cm + 2.8 cm + 3.5 cm + 4.2 cm + 4.9 cm + 5.6 cm + 6.3 cm + 7 cm
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Adjust space between Latin and Asian text, Adjust space between Asian text and numbers
Formatted: Adjust space between Latin and Asian text, Adjust space between Asian text and numbers
Formatted: Font: 10 pt
Formatted: Font: 10 pt
Formatted: Font: 10 pt
Formatted: Font: 10 pt

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

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

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

3.1
bit
atomic unit of information in a computer system

3.2
byte
sequence of 8 *bits* ~~(3.1)~~(3.1)

3.3
nibble
half a *byte* ~~(3.2),~~(3.2), which can be represented by a single hexadecimal digit

[SOURCE: ISO 20038:2017, 3.12]

3.4
data
ordered sequence of *bits* ~~(3.1)~~(3.1)

3.5
file
stored *data* ~~(3.4)~~(3.4) with a known number of *bits* ~~(3.1)~~(3.1) and a filename

3.6
stream
data ~~(3.4)~~(3.4) in transit with a known or unknown number of *bits* ~~(3.1)~~(3.1)

3.7
content
information organized to provide value to a user

3.8
digital content
manifestation of *content* ~~(3.7)~~(3.7) in form of *data* ~~(3.4)~~(3.4) structured according to a set of rules

3.9
metadata
~~*data* (3.4)~~*data* (3.4) that defines and describes other data

[SOURCE: ISO 24531:2013, 4.32]

3.10
seed metadata
initial *metadata* ~~(3.9)~~(3.9) used as input to a *hash algorithm* ~~(3.1)~~(3.1) function

iTeh Standards
(https://standards.iteh.ai)
Document Preview
ISO/PRF 24138
https://standards.iteh.ai/catalog/standards/iso/4bd4b8ba-b648-409e-8713-22cb3cbbe602/iso-prf-24138

© ISO 2024 – All rights reserved

Edited DIS - MUST BE USED FOR FINAL DRAFT

Formatted: Font: 11 pt, Bold
Formatted: Font: 11 pt, Bold
Formatted: Font: 11 pt, Bold
Formatted: Font: Bold
Formatted: HeaderCentered
Formatted: Adjust space between Latin and Asian text, Adjust space between Asian text and numbers, Tab stops: Not at 0.7 cm + 1.4 cm + 2.1 cm + 2.8 cm + 3.5 cm + 4.2 cm + 4.9 cm + 5.6 cm + 6.3 cm + 7 cm
Formatted: Adjust space between Latin and Asian text, Adjust space between Asian text and numbers
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: FooterPageRomanNumber

3.11
content format
set of rules used to structure *digital content* (3.8)(3.8)

3.12
media type
two-part *identifier* (3.15)(3.15) specifying the nature of the referenced *data* (3.4)(3.4)

[SOURCE: ISO/IEC 19757-4:2006, 3.9]

3.13
digital asset
file (3.5)(3.5) or *stream* (3.6)(3.6) encoded in conformance with a specific *content format* (3.11)(3.11)

3.14
referent
object which is identified

3.15
identifier
sequence of characters that uniquely denotes a *referent* (3.14)(3.14)

3.16
identifier system
system to enable the provision of *identifiers* (3.15)(3.15) for a given category of *referents* (3.14)(3.14)

3.17
content identifier
identifier (3.15)(3.15) whose *referent* (3.14)(3.14) is *content* (3.7)(3.7)

3.18
content-dependent identifier
content identifier (3.18)(3.18) whose *data* (3.4)(3.4) depends on the *digital content* (3.8)(3.8) that it identifies

3.19
content recognition system
system whose primary purpose is to recognise *digital content* (3.8)(3.8) on a granular level

3.20
algorithm
set of instructions

3.21
hash algorithm
deterministic *algorithm* (3.20)(3.20) that produces fixed-length *data* (3.4)(3.4) from an input of arbitrary-length data

3.22
hash digest
result of processing *data* (3.4)(3.4) with a *hash algorithm* (3.21)(3.21)

iTeh Standards
(https://standards.iteh.ai)
Document Preview
https://standards.iteh.ai/catalog/standards/iso/4bd4b8ba-b648-409e-8713-22cb3cbbe602/iso-prf-24138

© ISO 2024 – All rights reserved

Formatted: Font: 11 pt, Bold
Formatted: Font: 11 pt, Bold
Formatted: Font: Bold
Formatted: HeaderCentered, Left
Formatted: Font: 11 pt, Bold
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: FooterPageRomanNumber

3.23
cryptographic hash function
computationally efficient function mapping binary strings of arbitrary length to binary strings of fixed length, such that it is computationally infeasible to find two distinct values that hash into a common value

3.24
similarity hash
hash digest ~~(3.22)~~(3.22) that preserves correlations between inputs to the *hash algorithm* ~~(3.21)~~(3.21)

3.25
content defined chunking
CDC
method to split *data* ~~(3.4)~~(3.4) into variable length chunks based on internal features such that chunk boundaries are more resistant to *byte* ~~(3.2)~~(3.2) shifting

3.26
actor
human or non-human (hardware or software) entity that interacts with a system

3.27
~~merkle~~Merkle tree
tree data structure in which every leaf node is labelled with the *hash digest* ~~(3.22)~~(3.22) of a data element and every non-leaf node is labelled with the hash digest of the labels of its child nodes

3.28
~~merkle~~Merkle root
root node of a *~~merkle~~Merkle tree* ~~(3.27)~~(3.27)

[SOURCE: ISO 22739:~~2020~~2024, 3.~~46~~57]

3.29
ISCC processor
application that generates ISCCs for *digital content* ~~(3.8)~~(3.8)

3.30
plain text
~~*data* (3.4)~~*data* (3.4) with a known text encoding that can be transcoded to Unicode

3.31
whitespace
nondisplaying formatting characters such as spaces, tabs, etc., that are embedded within a block of free text

[SOURCE: ISO/IEC/IEEE 31320-2:2012, 3.1.210]

4 Structure and format of the ISCC

4.1 General structure

~~A. An ISCC shall be composed of an ISCC-HEADER and an ISCC-BODY (see Figure 1).~~

~~© ISO 2024 – All rights reserved~~

© ISO 2024 – All rights reserved

Edited DIS - MUST BE USED FOR FINAL DRAFT

iTeh Standards
(https://standards.iteh.ai)
Document Preview
ISO/PRF 24138
https://standards.iteh.ai/catalog/standards/iso/4bd4b8ba-b648-409e-8713-22cb3cbbe602/iso-prf-24138

Formatted: Font: 11 pt, Bold
Formatted: Font: 11 pt, Bold
Formatted: Font: 11 pt, Bold
Formatted: Font: Bold
Formatted: HeaderCentered
Formatted: Font: Not Italic
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Default Paragraph Font
Formatted: Adjust space between Latin and Asian text, Adjust space between Asian text and numbers, Tab stops: Not at 0.71 cm
Formatted: FooterPageRomanNumber