INTERNATIONAL STANDARD

ISO 18232

First edition
2006-04-01

Health Informatics — Messages and communication — Format of length limited globally unique string identifiers

Informatique de santé — Messages et communication — Format des identifiants uniques universels codés en caractère et limités en longueur

Reference number
ISO 18232:2006(E)

Contents

Page

Foreword iv
Introduction v
1 Scope 1
2 Normative references 1
3 Terms, definitions and abbreviated terms 1
3.1 Terms and definitions 1
3.2 Abbreviated terms 2
4 Globally unique string identifier field 2
5 ISO/IEC data object identifier GUSI exchange format 2
5.1 General 2
5.2 Logical data format 2
5.3 Character representation 3
5.4 Encoding 3
5.5 Integer component representation 3
5.6 Integer component separator 3
5.7 Interpretation of the data object identifier string 3
6 128 bit globally unique string identifier format 3
6.1 General 3
6.2 Logical data format 3
6.3 Character representation 3
6.4 UUID components 3
6.5 Encoding 3
Annex A (informative) Scenarios 4
Annex B (informative) Nature and application of Data Object Identifiers 5
Annex C (informative) ASN.1 specification for the exchange of OID and UUIDs 8

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.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO 18232 was prepared by Technical Committee ISO/TC 215, *Health informatics*.

Introduction

Data objects used in medicine include reports and results of diagnostic procedures which are stored and exchanged in electronic form, and objects such as templates. Applications must be able to find the location and the identification of such objects. Object identifiers are often numeric in form. This International Standard provides a means of exchanging globally unique identifiers expressed as character strings. It is not concerned with the specification of the location from which a data object may be retrieved.

A healthcare service for a patient is delivered in identifiable parts which may be termed healthcare service items. A healthcare service item can be performed for a patient by a healthcare professional, or a healthcare professional may request a healthcare service item to be performed by another healthcare professional or by a healthcare service department such as a medical imaging service department.

Healthcare service item results arise from numerical measurement or assessment by a healthcare professional. Individual numerical results may be included within report text, perhaps in a table. Sets of numerical results may be presented visually e.g. waveform (graph) or image (picture).

Results that consist of a large number of measured values such as a waveform or digital image are known as data objects.

To allow safe use in medicine, all data objects must be identified by a globally unique identifier (GUI), such as an ISO OID or binary MS GUI. The GUI allocated to a data object is attached to the data object (e.g. by including it within a computer file header section). The reference to a data object includes the GUI of the data object as well as the path to the data object. The application that retrieves the data object can verify that the correct data have been retrieved by matching the GUI in the reference to the GUI attached to the data object. See Annex A for relevant scenarios.

It may be noted that:

a) the issue of the location of a data object is separate from the issue of its identity; indeed several identical copies of the object may exist;

b) a globally unique data object identifier is intended for machine use and may be quite large;

c) a short, user-friendly, locally unique identifier is often required in addition to the globally unique identifier for human use. (This is outside the scope of this International Standard.)

Globally unique identifiers are already specified in various standards and the intention of this International Standard is to provide a specification for a common format for the exchange of commonly used globally unique identifiers expressed as alphanumeric strings.

A logical data format for globally unique identifiers constructed from a sequence of integers is defined by ISO/IEC 8824-1. Identifiers based on ISO/IEC 8824-1 are widely used in medical imaging. 128 bit universal unique identifiers (UUIDs) are widely used in the MS Windows environment. This International Standard specifies the format of alphanumeric string fields for the exchange of globally unique string identifiers (GUSI)

Health Informatics — Messages and communication — Format of length limited globally unique string identifiers

1 Scope

This International Standard specifies the encoding and length for globally unique identifiers for data objects used in healthcare exchanged as alphanumeric strings.

The technologies used for data storage, location and communication are outside the scope of this International Standard.

2 Normative references

The following referenced documents are indispensable for the application 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 646, *Information technology — ISO 7-bit coded character set for information interchange*

ISO/IEC 8824-1:2002, *Information technology — Abstract Syntax Notation One (ASN.1): Specification of basic notation — Part 1*

ISO/IEC 9834 (all parts), *Information Technology — Open Systems Interconnection — Procedures for the operation of OSI Registration Authorities*

DICOM PS 3.3, *Digital Imaging and Communications in Medicine — Part 3: Information Object Definitions*

DICOM PS 3.4, *Digital Imaging and Communications in Medicine — Part 4: Service Class Specifications*

DICOM PS 3.5, *Digital Imaging and Communications in Medicine — Part 5: Data Structures and Encoding*

3 Terms, definitions and abbreviated terms

3.1 Terms and definitions

3.1.1
information object
well-defined piece of information, definition, or specification which requires a name in order to identify its use in communication

NOTE Adapted from ISO/IEC 8824-1.

3.1.2
object identifier
OID
value (distinguishable from all other such values) that is associated with an information object

NOTE Adapted from ISO/IEC 8824-1.

3.1.3
data object identifier
sequence of integer components constructed as specified in ISO/IEC 8824-1, having a root sequence of components issued by a national standards body and sequence of following integer components whose uniqueness is guaranteed by the organization that defined it

NOTE Which is guaranteed to be globally unique by that national body registered as specified by ISO 9834.

3.1.4
unique identifier
identifier that is different from any other such identifier within a given namespace

3.1.5
globally unique identifier
identifier that is different from any other such identifier in any domain namespace

3.1.6
globally unique string identifier
alphanumeric string of maximum length 64 characters, which is different from any other alphanumeric string that has been or will be exchanged according to the provisions of this International Standard

3.2 Abbreviated terms

GUI globally unique identifier

OID object identifier

UID unique identifier (DICOM)

GUSI globally unique string identifier

UUID universal unique identifier

4 Globally unique string identifier field

When communicated by means of electronic data exchange the field containing the GUSI shall conform to the following:

a) length: 64 octets;

b) type: alphanumeric.

NOTE The object identifier field might be included as part of the data object metadata or might be part of a "wrapper" containing the data object. However this issue is outside the scope of this International Standard.

5 ISO/IEC data object identifier GUSI exchange format

5.1 General

Applications that conform to this clause shall support the use of object identifiers (OIDs) as specified by ISO/IEC 8824-1 and shall be constructed and encoded according to 5.2 to 5.7.

5.2 Logical data format

Data object identifiers shall be constructed from a sequence of integer components in accordance with the provisions of ISO/IEC 8824-1 and ISO/IEC 9834-7.

5.3 Character representation

Data object identifier integer components shall be represented as a character string using the following characters:

"0", "1", "2", "3", "4", "5", "6", "7", "8", "9", "."

NOTE See Annex A for an example.

5.4 Encoding

A data object identifier shall be encoded as an alphanumeric string of 8 bit octets as specified by ISO/IEC 646.

5.5 Integer component representation

An integer component shall not start with a zero character unless the component consists of a single zero.

5.6 Integer component separator

A single full stop (2E hexadecimal) character shall separate sequential components.

NOTE There is no leading or trailing full stop.

5.7 Interpretation of the data object identifier string

The root components allocated by a national body for use by an organization as the leading components of all data object identifiers created by that organization may carry meaning for the national body. However, applications shall not infer any semantic information from any component of the data object identifier.

6 128 bit globally unique string identifier format

6.1 General

The alphanumeric field specified in Clause 4 may also be used to hold a 128 bit binary number encoded using hexadecimal notation constructed as an alphanumeric string as specified in 6.2 to 6.5. See Annex C.3. Applications that conform to this International Standard and exchange 128 bit identifiers shall conform to the indications given in 6.2 to 6.5.

6.2 Logical data format

The logical data format shall be a 128 bit binary number.

6.3 Character representation

128 bit data object identifiers shall be represented by a sequence of octets using the following characters:

"0", "1", "2", "3", "4", "5", "6", "7", "8", "9", "a", "b", "c", "d", "e", "f", "-"

6.4 UUID components

The character string shall consist of a sequence of 5 components of lengths 8, 4, 4, 4 and 12 octets respectively, separated by a hyphen octet (2D hexadecimal)

6.5 Encoding

A data object identifier shall be encoded as an alphanumeric string of 8 bit octets as specified by ISO/IEC 646.

NOTE This clause supports the identifiers as given in the specifications of universal unique identifiers (UUIDs). See Annex B.

Annex A
(informative)

Scenarios

A.1 Verification of the identity of a referenced data object

A diagnostic procedure report may reference procedure products such as curves, waveforms, histograms and images recorded in electronic form. The observer of the results presented must be confident that the data presented are that which were selected by the clinician who created the report.

The appropriate use of OSI object identifiers as defined by ISO/IEC 8824-1 and ISO/IEC 9834-7 can enable the implementation of procedures to facilitate this aspect of security. For example, a report may include a pointer to a data object such as an image that comprises a unique identifier, the name of the file containing the image file and the path to the local copy of the file. When the application retrieves the file it can verify that the OID associated with the file is the same as the OID specified within the pointer. The UID associated with the file might be stored within a data header area or be incorporated in a "wrapper" containing the file.

A.2 Merging of patient records from different institutions

Records in electronic from are received by an institution and it is required to merge these with existing records. The availability of a globally unique identifier constitutes a unique database key field value, which facilitates the merging of data from different institutions.

A.3 Access to stored datasets within an institution

An institution implements an archive which contains the results of diagnostic procedures in electronic form stored in files and needs to implement a means to access them. A database is created that links the globally unique identifiers of the data objects to the path to the corresponding stored data files. This database also provides the basis for control of access according to a person's access rights.

A.4 Access to results referenced by reports that are sent outside the institution

A GP (General Practitioner) receives the results of a diagnostic procedure including a report and presented result files in encrypted form. The GP system does not support the use of UIDs as is done within the institution that produced the results. The data objects, once decrypted, contain embedded identification to permit verification of content by the use of appropriate software.

Annex B
(informative)

Nature and application of Data Object Identifiers

B.1 ISO/IEC object identifiers

An ISO OID as used in healthcare consists of a globally unique sequence of integer numbers. The first few numbers in the sequence, which are called “the root”, have been supplied to the organisation that created the OID by a national standards body. The issuing national standards body guarantees that the “root” supplied is globally unique. Subsequent, root-dependent, members of the sequence have been allocated by the organisation that requested the “root” by a method which ensures that identical sequences are never allocated. The details of this process are described in Clause 6 of ISO/IEC 9834-1:2005 in terms of a so called RH-name tree whose arcs correspond with members of the sequence of numbers, apart from the last member which is called a leaf. Although the standard was written to enable unique identifiers for computer use to be allocated to classes of object, ISO OIDS are also used in healthcare to identify instances of objects which often have also been allocated an identifier for human use.

It might be thought that an integer field of sufficient size to support a unique identifier for every data object to be created by an institution would be of impractical size. However, it should be borne in mind that, according to current estimates, 50 decimal digits are sufficient to enumerate all of the atoms in our planet. The justification for the choice of a 64 character field is given in B.2 below, which also suggests strategies for ensuring that identifiers generated by an application are indeed globally unique.

B.2 Example of ISO/IEC OID

The logical data structure is defined in ISO/IEC 8824-1.

The method to be used by national bodies is specified in ISO/IEC 9834-7. The following table is adapted from ISO/IEC 9834-1:1993.

Table B.1 — ISO/IEC OID using example values

RH-name: integer value	Object identifier
Joint-iso-ccit(2)	{ 2 }
Country(16)	{ 2 16 }
Country-name(840)	{ 2 16 840 }
State-or-province(46)	{ 2 16 840 46 }
Organization(3125)	{ 2 16 840 46 3125 }
Organisation-unit(3)	{ 2 16 840 46 3125 3 }
Organisation-unit-allocated-unique-object-number(1984675)	{ 2 16 840 46 3125 3 1984675 }

The penultimate component [Organisation-unit(3) in the Table B.1] is allocated to a department by the organization. The final component is allocated by the department (not part of the example as shown in ISO/IEC 9834-1:1993).

If the last (Organisation-unit-allocated-unique-object-number) integer component (1984675) is unique within the department then the alphanumeric string "2.16.840.46.3125.3.1984675" is a globally unique identifier because the leading sequence of 6 components is guaranteed to be unique by the national body that issued it.

Many systems create the last component by taking a departmentally unique system number and appending the date and time in seconds expressed as an integer. Some systems additionally append an integer component following the date/time sequence that is incremented each time a data object identifier is created and which is periodically reset to 1. This greatly reduces the possibility that an allocated final component is not unique within the department.

Table B.2 suggests possible maximum sizes for some typical components. In the very unlikely event that a system needed to create an OID that would not fit within the 63 character plus null field because too many identifiers had been already created, then a new root could be obtained or a new device number allocated.

Table B.2 — Component use

Component number	Component use	Characters
1-4	Root	12
5	One hundred million healthcare institutions	9
6	One hundred thousand departments per institution	6
7	One hundred thousand devices per department	6
8	Date and time to the nearest second	14
9	Incremented 3 digit integer	4
	Total chars to be stored in 64 character field	51
	12 chars remain unused out of 63 available – enough for one hundred thousand million more OIDs	

B.3 DICOM use of ISO OIDs

The DICOM standard includes the specification of globally unique identifiers based on the content of ISO/IEC 8824-1 and ISO/IEC 9834. OSI Object Identifiers (OIDs). ISO/IEC 9834-1:2005, Annex C, shows examples such as {2 16 840 46 3125 3}.

The first few components are supplied by a national standards body and identify it. The last but one component (3125) is allocated to the organization that requested the identifier. The last has been allocated by the organization to one of its departments. The department management must ensure the uniqueness of component(s) in the sequence following the root that has been allocated to it. The string comprising the root and subsequent components is a globally unique string.

DICOM specifies the use of a 64 ASCII (ISO/IEC 646) character numeric string data type for the exchange of ISO OID components and specifies a separator character as full stop (ASCII 46 decimal/2E Hex). The resulting object identifier is called a DICOM unique identifier (UID). Thus the example above would be stored as "2.16.840.46.3125.3". The internal representation of DICOM UIDs may be a binary number derived from the components or a character string. Modern hashing techniques applied in an alphanumeric string search can be almost as efficient as a binary number search.

B.4 UUID

A number of web-based applications use 128 bit unique identifiers UUIDs, represented by hexadecimal characters and the hyphen (-) when exchanged.

The following formal definition of the UUID used by several software vendors can be found at http://www.opengroup.org/onlinepubs/9629399/apdxa.htm.

B.5 String representation of UUIDS

For use in human readable text, a UUID string representation is specified as a sequence of fields, some of which are separated by single dashes.

Each field is treated as an integer and has its value printed as a zero-filled hexadecimal digit string with the most significant digit first. The hexadecimal values, a to f inclusive, are output as lower case characters, and are case insensitive on input. The sequence is the same as the UUID constructed type.

The formal definition of the UUID string representation is provided by the following extended BNF:

```
UUID                    = <time_low> <hyphen> <time_mid> <hyphen>
                            <time_high_and_version> <hyphen>
                            <clock_seq_and_reserved>
                            <clock_seq_low> <hyphen> <node>
time_low                  = <hexOctet> <hexOctet> <hexOctet> <hexOctet>
time_mid                  = <hexOctet> <hexOctet>
time_high_and_version     = <hexOctet> <hexOctet>
clock_seq_and_reserved    = <hexOctet>
clock_seq_low             = <hexOctet>
node                    = <hexOctet><hexOctet><hexOctet>
                            <hexOctet><hexOctet><hexOctet>
hexOctet                  = <hexDigit> <hexDigit>
hexDigit                = <digit> | <a> | <b> | <c> | <d> | <e> | <f>
digit                     = "0" | "1" | "2" | "3" | "4" | "5" | "6" | "7" |
                            "8" | "9"
hyphen                  = "-"
a                       = "a" | "A"
b                       = "b" | "B"
c                       = "c" | "C"
d                       = "d" | "D"
e                       = "e" | "E"
f                       = "f" | "F"
```

The following is an example of the string representation of a UUID:

2fac1234-31f8-11b4-a222-08002b34c003

Annex C
(informative)

ASN.1 specification for the exchange of OID and UUIDs

```
UUID-octets ::=  SEQUENCE
{
octet1          HexOctet,
octet2          HexOctet,
octet3          HexOctet,
octet4          HexOctet,
octet5          HexOctet,
octet6          HexOctet,
octet7          HexOctet,
octet8          HexOctet,
octet9          IA5String "-",
octet10   HexOctet,
octet11   HexOctet,
octet12   HexOctet,
octet13   HexOctet,
octet14   HexOctet,
octet15   HexOctet,
octet16   HexOctet,
octet17   HexOctet,
octet18   HexOctet,
octet19   IA5String "-",
octet20   HexOctet,
octet21   HexOctet,
octet22   HexOctet,
octet23   HexOctet,
octet24   IA5String "-",
octet25   HexOctet,
octet26   HexOctet,
octet27   HexOctet,
octet28   HexOctet,
octet29   HexOctet,
octet30   HexOctet,
octet31   HexOctet,
octet32   HexOctet,
octet33   HexOctet,
octet34   HexOctet,
octet35   HexOctet,
octet36   HexOctet
}
HexOctet              ::= CHOICE
{
Digit1-9,
d0   IA5String       "0",
a    IA5string"a" ,
b    IA5string"b" ,
c    IA5string"c" ,
d    IA5string"d" ,
e    IA5string"e" ,
f    IA5string"f"
}
Digit ::= CHOICE
```

```
{
IA5String      "0",
Digit1-9
}
Digit1-9                                                    ::=CHOICE
{
d1   IA5String      "1"                                             ,
d2   IA5String      "2"                                             ,
d3   IA5String      "3"                                             ,
d4   IA5String      "4"                                             ,
d5   IA5String      "5"                                             ,
d6   IA5String      "6"                                             ,
d7   IA5String      "7"                                             ,
d8   IA5String      "8" ,
d9   IA5String      "9"
}
OID-Octets ::= SEQUENCE
{
initial-component SEQUENCE OF
    {
    OID-component,
    IA5String ".   "
    }
final-component
{
OID-component SEQUENCE
{
CHOICE
    {
    IA5String "0",
    SEQUENCE
        {
        first-digit      Digit1-9,       -- First octet of a non-zero component may not be zero
        CHOICE
        {
        NULL,
        other-digits   SEQUENCE OF Digit
        }
        }
    }
}
```

ICS 35.240.80

Price based on 9 pages