

Mapping DICOM Data Types to HL7 Data Types Proposal

Author: Flora Gilboa

IBM Haifa Research Lab

Revision: 0.1

Last update Date: 17th, September 2001.

Rationale:

In our work on CDA Level 3 proposal for DICOM Series, we met the demand to derive elements in CDA from the HL7 RIM. The RIM, as for today, holds a very small part of classes suitable for DICOM, e.g. Diagnostic_image class holds one attribute called subject_orientation_cd. Entering some more of DICOM attributes into HL7 can be done through R-MIMs (which have to maintain HL7 data types). Harmonizing DICOM into HL7, requires mapping DICOM data Types into HL7 Data Types. I hope this first version will be a trigger for discussion.

Comment:

As HL7 Data Types are on ballot now, this mapping should be updated once the HL7 data types will be changed.

Statement:

In the procedure of mapping x to y, there is always a compromise. In this case, DICOM types are naturally more specific than HL7 types, which are common to all medical areas. Therefore, while mapping specific to general data types, we lose information. In addition, DICOM defines Value Representation (VR) of a data element. VR describes the data type and format of the data element(s) value(s). For example, lets look at DICOM AS, Age String. DICOM defines: "A string of characters with one of the following formats -- nnnD, nnnW, nnnM, nnnY; where nnn shall contain the number of days for D, weeks for W, months for M, or years for Y. Example: "018M" would represent an age of 18 months. ". Mapping it to HL7 ST, Character String, losses this format.

Missing Type:

I have not found a parallel HL7 type for DICOM TM (Time) type. The HL7 TS, point in time, specifies a full date (yyyymmddhhmmss.uuuu) while TM is just the hhmmss.uuuu part. DICOM DA (Date) has the format of yyyymmdd and can be mapped to TS type by not supplying the hhmmss.uuuu part. (It is valid to omit the right hand side of TS, but not the left hand side).

Table 1: Mapping DICOM VRs to HL7 Data Types.

DICOM Data Types		HL7 Data Types		Comments
VR *	Full Name	Symbol	Full Name	
AE	Application Entity	ST	Character String	
AS	Age String	ST	Character String	
AT	Attribute Tag **	SET	Set	Need a conversion function
CS	Code String	ST	Character String	
DA	Date	TS	Point in Time	By omitting the hhmss.uuuu part
DS	Decimal String	REAL	Real Number	Need a conversion function
DT	Date Time	TS	Point in Time	
FL	Floating Point Single	REAL	Real Number	
FD	Floating Point Double	REAL	Real Number	
IS	Integer String	INT	Integer Number	Need a conversion function
LO	Long String	ST	Character String	
LT	Long Text	ST	Character String	
OB	Other Byte String **	ST	Character String	Not good enough
OW	Other Word String **	ST	Character String	Not good enough
PN	Person Name	PN	Person Name	
SH	Short String	ST	Character String	
SL	Signed Long	INT	Integer Number	
SQ	Sequence of Items	LIST	Sequence	
SS	Signed Short	INT	Integer Number	
ST	Short Text	ST	Character String	
TM	Time	?		
UI	Unique Identifier (UID)	II	Instance Identifier	
UL	Unsigned Long	INT	Integer Number	
UN	Unknown **	ST	Character String	Not good enough
US	Unsigned Short	INT	Integer Number	
UT	Unlimited Text	ST	Character String	

* VR - Value Representation.

** Look at Decision to be made section

Decisions to be made:

1. DICOM AT, Attribute tag, is an ordered pair of 16-bit unsigned integers, (gggg, eeee), where gggg equates to the Group Number and eeee equates to the Element Number within that Group. The data is represented in hexadecimal notation e.g. (0018,00FF).

Mapping possibilities:

- SET containing two INT, group number and element number.
- INT – by concatenation of the two numbers – ggggeeee.

In this mapping (table 1), I propose a SET of two INT as I think it is closer to DICOM intention. In both cases it seems that an application will need a conversion function.

2. DICOM OW, Other Word String & OB, Other Byte String – These VRs are string of words/bytes where the encoding of the contents is specified by the

negotiated Transfer Syntax (part of DICOM protocol). Mapping them to HL7 ST is not good enough and should be discussed more.

3. DICOM UN, Unknown – This VR is a string of bytes where the encoding of the contents is unknown. I mapped it to ST, but will be glad to get more inputs on this VR from DICOM people.

References:

1. DICOM Standard 2000 Final Draft, PS 3.5 Sections 6.2 & 7.5.
2. HL7 Data Types Part I,
http://www.hl7.org/home/content/HTML_menus/DataTypes1.htm#TS