

DICOM Correction Proposal

STATUS	Final Text
Date of Last Update	2014/01/05
Person Assigned	Andrei Leontiev (andrei.leontiev@ge.com)
Submitter Name	Harry Solomon
Submission Date	2013/05/27

Correction Number	CP-1311
Log Summary: Describe matching of wildcard characters encoded in string VRs	
Name of Standard PS 3.3 2011 PS 3.5 2011	
Rationale for Correction Use of the characters * and ? in attributes used for matching (especially Patient ID and Patient Name) conflicts with its use in query matching strings. In a query string there is no way to distinguish between use of these characters as literals for single value matching and use as wildcards for wildcard matching. For example, if Patient ID would be allowed to have * or ? characters, an MWL or UPS SCP application may reject a single value match string because it looks like an illegal wildcard match string, or a Patient Root C-FIND SCP may reject a STUDY, SERIES, or IMAGE level operation because the Patient ID Unique Key looks like a wildcard match string. Similarly, the attempt to perform single-value match for an Accession Number with one of these characters will be interpreted by an SCP as a wildcard match. The problem is most pronounced for the Attributes with VRs of AE, PN, SH and LO that are defined as Unique or Required Matching keys for Query/Retrieve, MWL and UPS services. To address the issue, explanatory notes are added to Section C.2.2.4 Wildcard Matching.	
Correction Wording:	

Amend PS 3.4 with the notes about use of wildcard matching in attributes with certain VRs

C.2.2.2.1 Single Value Matching

If the value specified for a Key Attribute in a request is non-zero length and if it is:

- a) not a date or time or datetime, contains no wild card characters
- b) a date or time or datetime, contains a single date or time or datetime with no “-“

then single value matching shall be performed. Except for Attributes with a PN Value Representation, only entities with values which match exactly the value specified in the request shall match. This matching is case-sensitive, i.e., sensitive to the exact encoding of the key attribute value in character sets where a letter may have multiple encodings (e.g., based on its case, its position in a word, or whether it is accented).

For Attributes with a PN Value Representation (e.g., Patient Name (0010,0010)), an application may perform literal matching that is either case-sensitive, or that is insensitive to some or all aspects of case, position, accent, or other character encoding variants.

Notes: 1. For multi-component names, the component group delimiter “=” (3DH) may be present in the Key Attribute value, but may give unexpected results if the SCP does not support matching on separate components but interprets the entire value literally as a single string. E.g., “Wang^XiaoDong=王^小東”

may or may not match “Wang^XiaoDong” or “王^小東”; wildcard matching without the component group delimiter, such as “*Wang^XiaoDong*” or “*王^小東*” may be necessary.

2. Using attributes with VR of AE, LO, PN and SH as matching keys will not allow single value matching on values that contain characters “*” and “?” – such queries will always be treated as queries with wildcard matching.

3. Attributes with VR of ST, LT and UT are intended for conveying narrative text and may contain wildcard characters “*” and “?”. Attempts to match on a string explicitly containing “*” or “?” will be treated as wildcard matching and thus may return multiple results rather than a single one.

C.2.2.2.4 Wild Card Matching

If the Attribute is not a date, time, signed long, signed short, unsigned short, unsigned long, floating point single, floating point double, other byte string, other word string, unknown, attribute tag, decimal string, integer string, age string or UID and the value specified in the request contains any occurrence of an “*” or a “?”, then “*” shall match any sequence of characters (including a zero length value) and “?” shall match any single character. This matching is case sensitive, except for Attributes with an PN Value Representation (e.g., Patient Name (0010,0010)).

For attributes with a PN value representation, including the case of extended negotiation of fuzzy semantic matching, wild card matching is implementation dependent and shall be specified in the conformance statement.

- Notes:
1. Wild card matching on a value of “*” is equivalent to universal matching.
 2. The wild card matching method specified by DICOM might not be supported by some non-DICOM multi-byte character text processors.
 3. For multi-component group names, the component group delimiter “=” (3DH) may be present in the Key Attribute value, but may give unexpected results if the SCP does not support matching on separate components but interprets the entire value literally. E.g., “*=*” or “*==*” may or may not return all strings, and hence is not equivalent to “*”, nor to universal matching.
- 4. Using attributes with VR of AE, LO, PN and SH as matching keys will not allow single value matching on values that contain characters “*” and “?” – such queries will always be treated as queries with wildcard matching.**
- 5. Attributes with VR of ST, LT and UT are intended for conveying narrative text and may contain wildcard characters “*” and “?”. Attempts to match on a string explicitly containing “*” or “?” will be treated as wildcard matching and thus may return multiple results rather than a single one.**