Jim,

This is going to be a little more complicated than you were hoping.  There are three spreadsheets attached.

The first (ListOfPrivateElementsWithDispositions.xls) is a list of every private tag we have seen in our PHI scans.  It also preserves the position in the data hierarchy in which the tag occurred.

In the database from which this is extracted we also maintain a record of all of the values we have seen in these tags.  This sometimes has multiple rows for a given private tag when we have seen a private tag with multiple values for VR.  For example:

|  |  |  |  |
| --- | --- | --- | --- |
| (group # , Creator Token, element #) | VR | Disposition | Name |
| (0019,"SIEMENS MR HEADER",13) | UN | k | Ima Abs Table Position |
| (0019,"SIEMENS MR HEADER",13) | SL | k | Ima Abs Table Position |
| (0019,"SIEMENS MR HEADER",13) | OB | k | Ima Abs Table Position |

Some implementations encode unknown private tags as ‘UN’ VR (proper DICOM when converting from default xfer\_syntax), sometimes they encode it as ‘OB’ (wrong, but apparently popular), and some implementations know the proper VR.

It has a column for disposition:

 “d” - delete

“o” - offset date

“oi” - offset date (epoch)

“k” - keep

“h” - hash did

“na” - (in sequence which was deleted)

The next (ListOfPrivateElementsFromDD.xls) is a list from our private tag data dictionary.  This dictionary is derived from the following sources:

1. dicom3tools
2. dcmtk
3. gdcm
4. The Wustl private tag knowledge base

This database is used in deriving names when scanning for PHI.

Not all of the tags in we have encountered in scanning have been incorporated into the DD yet (tags with <undef> in the name chain column in particular).  (For example, the owner '"IPKCREP 2Q00+0C  TMIGA ERPVITA EADAT” is a mangling of "PICKER P@2000+ CT IMAGE PRIVATE DATA”, which we haven’t bothered to define)

Tags with an owner like “Unnamed Private Block - xx” in the first spreadsheet are tags which were illegal in dicom (there was no tag defining the owner of the private block)

Speaking of private blocks, the second spreadsheet accommodates another illegal DICOM practice (having identical private tags with different definitions depending on which block is selected for encoding.  These definitions have 4-digit element numbers.  The rest have 2-digit element numbers.

Finally, some of the databases we sourced defined some repeating groups.  These are in the third spreadsheet.

I’m sure you are going to have some questions.  Feel free to ask.

Thanks,

Bill B