Canonical Encoding for DICOM

# General

1. Remove all Group Length elements, except FMI Group Length (0002,0000).
2. Remove all *Undefined Length* fields.
3. Convert *Implicit VR* to *Explicit VR*
4. All Datasets should have a Patient Identity Removed (0012,0062) element with appropriate value.
5. Pixel Data
   1. Remove all fragments
   2. Add basic offset table if multiple frames are present
   3. Add the Number of Frames (0014,3012) element if not present.
6. Convert UN to Correct VR whenever possible.
7. Replace retired elements with their equivalent if known.

# Private Data

1. Remove Private Group Length elements.
2. Remove any Illegal Private Group elements, i.e. (gggg, 0001) to (gggg,0009).
3. Add Private Creator for Private Data Elements without a creator. For each Private Data Subgroup add a created with token 'Missing Private Creator #X', where X increases monotonically.

Note:

1. Record all Private Elements in a private element database.

# De-Identification and Profiling

1. Add the Patient Identity Removed (0012,0062) attribute to the Root Dataset
   1. Add the De-Identification Method Code Sequence (0012,0064).
   2. Add the De-Identification Method (0012,0063), TODO: create a list of well-defined methods or profiles.
2. Add new FMI describing the de-identification application:
   1. TODO: describe the various values.
   2. TODO: get a DICOM UID branch for ACR.
3. All Date elements should be *normalized* to a patient or trial related value, or should be removed. In either case the Longitudinal Temporal Information Modified (0028,0303) element should be added to the Dataset with a value of "UNMODIFIED", "MODIFIED", or "REMOVED".
4. All DateTime elements should have the date-part *normalized* to a patient or trial related value, or should be removed.
5. All Time elements can be left as is (TODO: true?)

## Additional Attributes that Should be Added

## Additional Attributes that Might be Added

1. Patient's Age (0010,1010)
2. Encrypted Attributes Sequence (0400,0500)
   1. Create Encrypted Attributes Dataset and Modified Attribute Sequence (0400,0550)
   2. Security policy
   3. Key management policy
3. Create one or more instances of the Encrypted Attributes Data Set and copy Attributes to be protected into the (single) item of the Modified Attributes Sequence (0400,0550) of one or more of the Encrypted Attributes Data Set instances.