pydicom tutorial

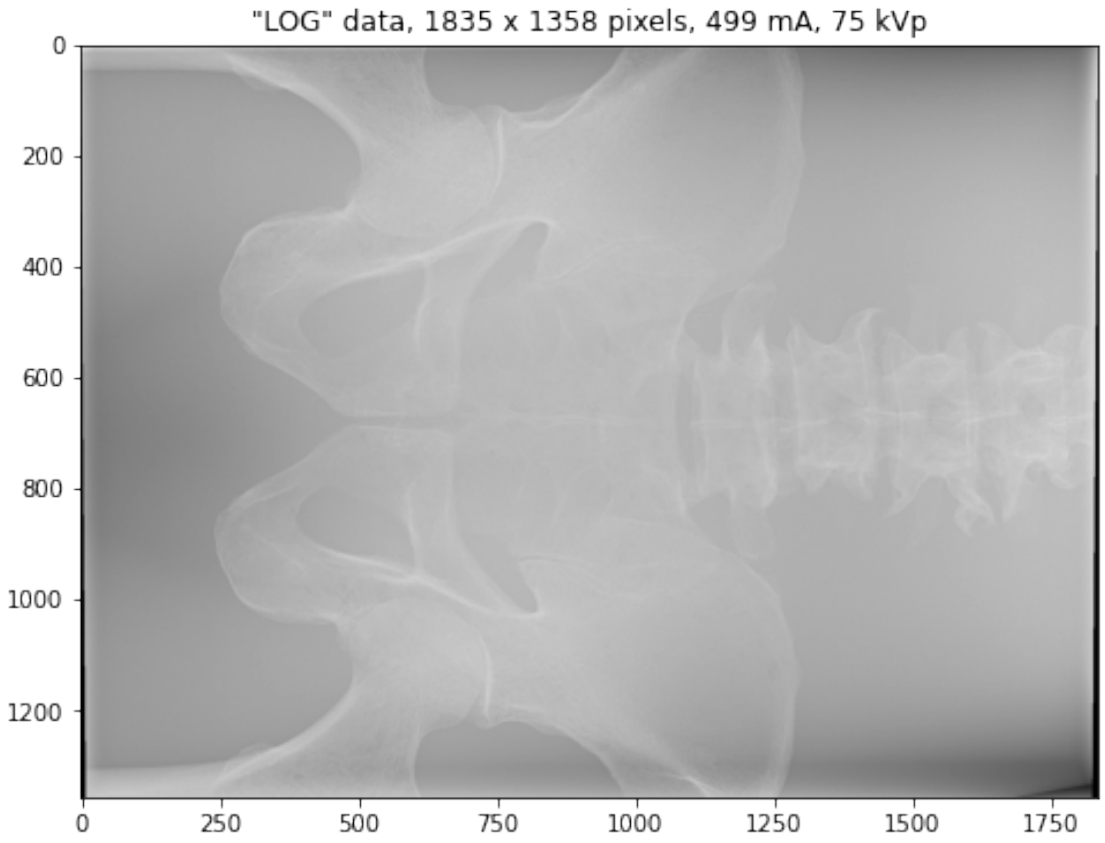
January 27, 2022

[1]:
```
import pydicom
import matplotlib.pyplot as plt
import os
```

We can use the pydicom function `dcmread` to load a dataset as a pydicom-specific class. Printing the dataset will display all its available attributes, corresponding to Dicom header tags. These may be used to access the data. For example, the image itself is saved as the `pixel_array` attribute. It is read in as a 2D numpy array of unsigned int16 values. Most other attributes are read in as strings.

For reference, here is a basic example.

[2]:
```
# first, read in the data and get a list of files
imd = 'imgs/'
L = sorted(os.listdir(imd))

# read the first dataset
i = 0
ds = pydicom.dcmread(imd+L[i])

# extract some key Dicom information
M = ds.pixel_array                          # the image
rows, cols = ds.Rows, ds.Columns            # image dimensions
pir = ds.PixelIntensityRelationship         # log or linear
mA = ds.XRayTubeCurrent
kVp = int(ds.KVP)

# plot the image
fig,ax = plt.subplots(1,1,figsize=[8,6])
ax.imshow(M, cmap='gray')
ax.set_title(f'"{pir}" data, {cols} x {rows} pixels, {mA} mA, {kVp} kVp')
plt.show()
```



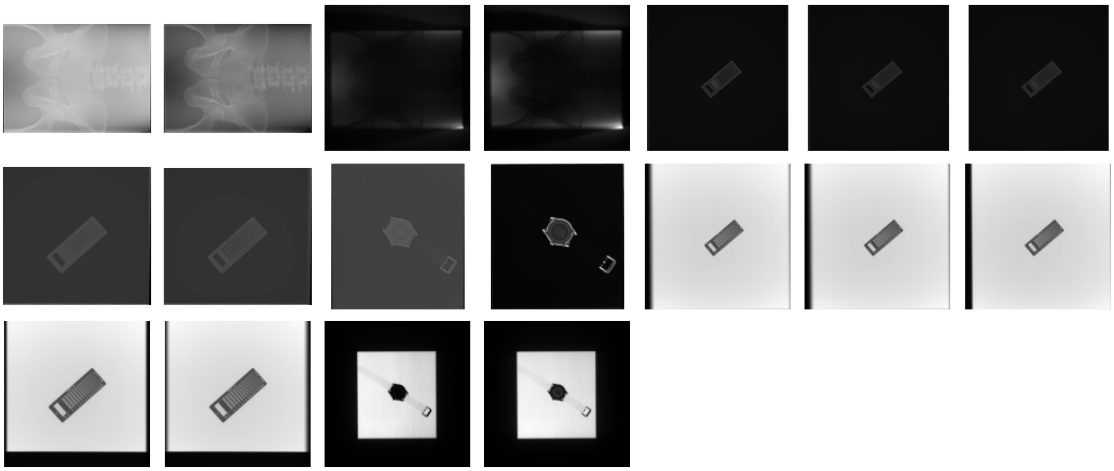
This is a image of the anthropomorphic phantom we scanned. The "LOG" indicates the pixel values are not raw and have been through some post-processing. There are also header attributes for the image size, tube current, and kVp.

Let's confirm that all the images look reasonable and check whether they use raw or log pixel values.

[3]:
```
yd, xd = 3,7
fig,ax = plt.subplots(yd,xd,figsize=[xd*2, yd*2])

for i in range(len(L)):
    ds = pydicom.dcmread(imd+L[i])

    # plot image
    ax[i//xd, i%xd].imshow(ds.pixel_array, cmap='gray')

    # print pixel relationship
    print(ds.PixelIntensityRelationship)

for i in range(xd*yd):
    # turn off axes so images look niver
    ax[i//xd, i%xd].axis('off')
```

```
fig.tight_layout()
plt.show()
```

```
LOG
LOG
LIN
LIN
LOG
LOG
LOG
LOG
LOG
LOG
LOG
LIN
LIN
LIN
LIN
LIN
LIN
LIN
```



This looks like all of our images, but some of the data is linear and some is logged. It is important to be mindful of this for our computations.

We can print out the entire Dicom header to see all available information.

[4]:
```
print(ds)
```

```
Dataset.file_meta -------------------------------
(0002, 0000) File Meta Information Group Length  UL: 168
```

```
(0002, 0001) File Meta Information Version       OB: b'\x00\x01'
(0002, 0002) Media Storage SOP Class UID         UI: Digital X-Ray Image Storage
- For Processing
(0002, 0003) Media Storage SOP Instance UID      UI:
1.2.840.113619.2.203.4.2147483647.1643241378.878071
(0002, 0010) Transfer Syntax UID                 UI: Explicit VR Little Endian
(0002, 0012) Implementation Class UID            UI: 1.2.840.113619.6.203
-------------------------------------------------
(0008, 0005) Specific Character Set              CS: 'ISO_IR 100'
(0008, 0008) Image Type                          CS: ['ORIGINAL', 'PRIMARY', '']
(0008, 0016) SOP Class UID                       UI: Digital X-Ray Image Storage
- For Processing
(0008, 0018) SOP Instance UID                    UI:
1.2.840.113619.2.203.4.2147483647.1643241378.878071
(0008, 0020) Study Date                          DA: '20220126'
(0008, 0021) Series Date                         DA: '20220126'
(0008, 0022) Acquisition Date                    DA: '20220126'
(0008, 0023) Content Date                        DA: '20220126'
(0008, 002a) Acquisition DateTime                DT: '20220126175618.000000'
(0008, 0030) Study Time                          TM: '161150.000'
(0008, 0031) Series Time                         TM: '161408.000'
(0008, 0032) Acquisition Time                    TM: '175618.000'
(0008, 0033) Content Time                        TM: '175618.000'
(0008, 0050) Accession Number                    SH: ''
(0008, 0060) Modality                            CS: 'DX'
(0008, 0068) Presentation Intent Type            CS: 'FOR PROCESSING'
(0008, 0070) Manufacturer                        LO: 'GE Healthcare'
(0008, 0080) Institution Name                    LO: 'UNIV. OF CHICAGO'
(0008, 0081) Institution Address                 ST: '5835 S. COTTAGE
GROVE\r\nMITCHELL TC114\r\nCHICAGO\r\nIL\r\n60637'
(0008, 0090) Referring Physician's Name          PN: ''
(0008, 1010) Station Name                        SH: 'MITTC114'
(0008, 1030) Study Description                   LO: ''
(0008, 103e) Series Description                  LO: 'Thumb'
(0008, 1040) Institutional Department Name       LO: 'RADIOLOGY'
(0008, 1050) Performing Physician's Name         PN: ''
(0008, 1070) Operators' Name                     PN: ''
(0008, 1090) Manufacturer's Model Name           LO: 'Discovery XR656'
(0008, 2218)  Anatomic Region Sequence  1 item(s) ----
   (0008, 0100) Code Value                          SH: 'T-D8810'
   (0008, 0102) Coding Scheme Designator            SH: 'SNM3'
   (0008, 0104) Code Meaning                        LO: 'Thumb'
   ---------
(0010, 0010) Patient's Name                      PN: 'Radiography Lab'
(0010, 0020) Patient ID                          LO: 'MPQC01262021'
(0010, 0030) Patient's Birth Date                DA: ''
(0010, 0032) Patient's Birth Time                TM: ''
(0010, 0040) Patient's Sex                       CS: 'O'
```

```
(0010, 1010) Patient's Age                          AS: ''
(0011, 0010) Private Creator                        LO: 'GEMS_GDXE_FALCON_04'
(0011, 1003) [Processed Series UID]                 UI:
1.2.840.113619.2.203.4.2147483647.1643235248.198669
(0011, 1004) [Acquisition Type]                     CS: 'SE'
(0011, 1005) [Acquisition UID]                      UI:
1.2.840.113619.2.203.4.2147483647.1643241375.209945
(0011, 1006) [Image Dose]                           DS: '0.372645'
(0011, 1009) [Non-Digital Exposures]                SL: 20
(0011, 1010) [Total Exposures]                      SL: 27
(0011, 1012) [Patient Size String]                  LT: 'MEDIUM_ADULT'
(0011, 1013) [SPS UID]                              UI:
1.2.840.113619.2.203.4.2147483647.1643235110.550923
(0011, 1015) [Detector ARC Gain]                    DS: '4400.0'
(0011, 1016) [Processing Debug Info]                LT: '1100,\n{1634 1559 414 1569
414 469 1634 479 }(8
elements)\r\n1200,\r\n1201,\r\n2100,\r\n2500,\r\n2501,\r\n2502,\r\n'
(0011, 1017) [Override mode]                        CS: 'DISABLE'
(0011, 1019) [Film Speed Selection]                 DS: '100.0'
(0011, 1031) [Detected Field of View]               IS: [1632, 1547, 419, 1557,
419, 449, 1632, 459]
(0011, 1034) [Compensated Detector Exposure]        DS: '55.839879'
(0011, 1036) [Median Anatomy Count Value]           DS: '12027.0'
(0011, 1037) [DEI lower and upper limit values]     DS: [-2.000000, -1.000000,
1.000000, 2.000000]
(0011, 1042) [View IP]                              LO: 'antero-posterior'
(0011, 1044) [Receptor Type]                        CS: 'TABLE'
(0011, 1047) [Unknown]                              DS: '810.058228'
(0011, 1059) Private tag data                       CS: 'NORMAL'
(0011, 1060) Private tag data                       CS: 'HORIZONTAL'
(0011, 1064) Private tag data                       CS: 'URP_DETECTOR'
(0011, 1065) Private tag data                       LO: 'GlobalURPTable'
(0011, 1066) Private tag data                       SH: 'No'
(0011, 1067) Private tag data                       SH: 'Front'
(0011, 1068) Private tag data                       LO: '40:f4:a0:0:c:4c'
(0011, 1069) Private tag data                       DS: '-1.0'
(0011, 106d) Private tag data                       DS: '218.5712'
(0011, 1076) Private tag data                       CS: ['R', 'H']
(0011, 1077) Private tag data                       CS: 'YES'
(0011, 1078) Private tag data                       DS: '128.0'
(0011, 1079) Private tag data                       DS: '14152.05874'
(0011, 107a) Private tag data                       SS: 8
(0011, 107b) Private tag data                       SS: 8
(0011, 107c) Private tag data                       CS: 'YES'
(0011, 107e) Private tag data                       DS: '180.0'
(0011, 1080) Private tag data                       LT: '2'
(0011, 1081) Private tag data                       LO: 'AP'
(0011, 1082) Private tag data                       LO: 'GRID_FREQ_70_LP_CM'
```

```
(0018, 0015) Body Part Examined                 CS: 'THUMB'
(0018, 0060) KVP                                DS: '125.0'
(0018, 1020) Software Versions                  LO:
'dm_Platform_release-V14_FW48-4'
(0018, 1030) Protocol Name                      LO: 'THUMB_RIGHT'
(0018, 1110) Distance Source to Detector        DS: '903.0'
(0018, 1111) Distance Source to Patient         DS: '829.0'
(0018, 1147) Field of View Shape                CS: 'RECTANGLE'
(0018, 1149) Field of View Dimension(s)         IS: [243, 222]
(0018, 1150) Exposure Time                      IS: '10'
(0018, 1151) X-Ray Tube Current                 IS: '100'
(0018, 1152) Exposure                           IS: '1'
(0018, 1153) Exposure in uAs                    IS: '1008'
(0018, 115e) Image and Fluoroscopy Area Dose Pro DS: '0.812122'
(0018, 1160) Filter Type                        SH: 'NONE'
(0018, 1164) Imager Pixel Spacing               DS: [0.183610, 0.183610]
(0018, 1166) Grid                               CS: ['FIXED', 'FOCUSED']
(0018, 1190) Focal Spot(s)                      DS: '0.6'
(0018, 1405) Relative X-Ray Exposure            IS: '372'
(0018, 1411) Exposure Index                     DS: '5887.733905'
(0018, 1412) Target Exposure Index              DS: '814.00001'
(0018, 1413) Deviation Index                    DS: '8.6'
(0018, 1508) Positioner Type                    CS: ''
(0018, 1700) Collimator Shape                   CS: 'POLYGONAL'
(0018, 1720) Vertices of the Polygonal Collimato IS: [1621, 1546, 401, 1556,
401, 456, 1621, 466]
(0018, 5101) View Position                      CS: 'AP'
(0018, 6000) Sensitivity                        DS: '215.383703'
(0018, 7000) Detector Conditions Nominal Flag   CS: 'YES'
(0018, 7001) Detector Temperature               DS: '27.700001'
(0018, 7004) Detector Type                      CS: 'SCINTILLATOR'
(0018, 7005) Detector Configuration             CS: 'AREA'
(0018, 700a) Detector ID                        SH: 'UA45725-2'
(0018, 701a) Detector Binning                   DS: [1, 1]
(0018, 7020) Detector Element Physical Size     DS: [0.200000, 0.200000]
(0018, 7022) Detector Element Spacing           DS: [0.200000, 0.200000]
(0018, 7024) Detector Active Shape              CS: 'RECTANGLE'
(0018, 7026) Detector Active Dimension(s)       DS: [409.600000, 409.600000]
(0018, 7030) Field of View Origin               DS: [0.000000, 0.000000]
(0018, 7032) Field of View Rotation             DS: '0.0'
(0018, 7034) Field of View Horizontal Flip      CS: 'NO'
(0018, 704c) Grid Focal Distance                DS: '1000.0'
(0018, 7050) Filter Material                    CS: 'COPPER'
(0018, 7052) Filter Thickness Minimum           DS: '0.0'
(0018, 7054) Filter Thickness Maximum           DS: '0.0'
(0018, 7060) Exposure Control Mode              CS: 'MANUAL'
(0018, 7062) Exposure Control Mode Description  LT: ''
(0020, 000d) Study Instance UID                 UI:
```

```
1.2.840.113619.2.203.4.2147483647.1643235110.551096
(0020, 000e) Series Instance UID                UI:
1.2.840.113619.2.203.4.2147483647.1643235248.198658
(0020, 0010) Study ID                           SH: 'Def_Study_ID'
(0020, 0011) Series Number                      IS: '1629'
(0020, 0013) Instance Number                    IS: '7'
(0020, 0020) Patient Orientation                CS: ['R', 'H']
(0020, 0062) Image Laterality                   CS: 'R'
(0028, 0002) Samples per Pixel                  US: 1
(0028, 0004) Photometric Interpretation         CS: 'MONOCHROME1'
(0028, 0010) Rows                               US: 2022
(0028, 0011) Columns                            US: 2022
(0028, 0100) Bits Allocated                     US: 16
(0028, 0101) Bits Stored                        US: 14
(0028, 0102) High Bit                           US: 13
(0028, 0103) Pixel Representation               US: 0
(0028, 0300) Quality Control Image              CS: 'NO'
(0028, 0301) Burned In Annotation               CS: 'NO'
(0028, 1040) Pixel Intensity Relationship       CS: 'LIN'
(0028, 1041) Pixel Intensity Relationship Sign  SS: 1
(0028, 1052) Rescale Intercept                  DS: '0.0'
(0028, 1053) Rescale Slope                      DS: '1.0'
(0028, 1054) Rescale Type                       LO: 'US'
(0028, 2110) Lossy Image Compression            CS: '00'
(0032, 1060) Requested Procedure Description    LO: ''
(0040, 0244) Performed Procedure Step Start Date DA: '20220126'
(0040, 0245) Performed Procedure Step Start Time TM: '161150.000'
(0040, 0253) Performed Procedure Step ID        SH: '164323511090911'
(0040, 0254) Performed Procedure Step Descriptio LO: 'Performed Desc'
(0040, 0302) Entrance Dose                      US: 0
(0040, 0310) Comments on Radiation Dose         ST: '125%'
(0040, 0555)  Acquisition Context Sequence  0 item(s) ----
(0040, 8302) Entrance Dose in mGy               DS: '0.372645'
(0045, 0010) Private Creator                    LO: 'GEMS_FALCON_03'
(0045, 1062) [User Window Center]               IS: '7140'
(0045, 1063) [User Window Width]                IS: '14024'
(0045, 1065) [Requested Detector Entrance Dose]  IS: '14250'
(0045, 1069) [Collimator rotation]              IS: '0'
(0045, 1072) [Collimator Width]                 IS: '218'
(0045, 1073) [Collimator Height]                IS: '244'
(0054, 0220)  View Code Sequence  1 item(s) ----
   (0008, 0100) Code Value                         SH: 'R-10206'
   (0008, 0102) Coding Scheme Designator           SH: 'SNM3'
   (0008, 0104) Code Meaning                       LO: 'antero-posterior'
   ---------
(0088, 0200)  Icon Image Sequence  1 item(s) ----
   (0028, 0002) Samples per Pixel                  US: 1
   (0028, 0004) Photometric Interpretation         CS: 'MONOCHROME1'
```

```
   (0028, 0010) Rows                                US: 64
   (0028, 0011) Columns                             US: 64
   (0028, 0100) Bits Allocated                      US: 8
   (0028, 0101) Bits Stored                         US: 8
   (0028, 0102) High Bit                            US: 7
   (0028, 0103) Pixel Representation                US: 0
   (7fe0, 0010) Pixel Data                          OB: Array of 4096 elements
   ---------
(2050, 0020) Presentation LUT Shape             CS: 'INVERSE'
(7fdf, 0010) Private Creator                    LO:
'GEMS_GDXE_ATHENAV2_INTERNAL_USE'
(7fdf, 1010) [PPS Stream]                       LT: 'PPSTREAM|XRPP:|2.0|FILE|/d
atabase/PixelData///1.2.840.113619.2.203.4.2147483647.1643235110.551096/1.2.840.
113619.2.203.4.2147483647.1643235248.198658/1.2.840.113619.2.203.4.2147483647.16
43241378.878071'
(7fdf, 1011) [Pixel Data References (temporary)] LT:
'PPSTREAM|XRPP:|2.0|FILE|/database/PixelData//DefaultImageLocation-8366199105'
(7fdf, 1025) [PPS Status]                       CS: 'COMPLETED'
(7fe0, 0000) Group Length                       UL: 8176980
(7fe0, 0010) Pixel Data                         OW: Array of 8176968 elements
```

[ ]: