1. The sources of data will be images and textual in nature
   1. Filetypes ending in .dcm or .pac
   2. Data will come from the Cancer Imaging Archive by NIH’s National Cancer Institute
   3. The data will be within the Gigabyte range
      1. from 10GB to 200GB depending on the image
   4. The data represents the patient and its status
2. Requirements for Processing and Analysis
   1. Pydicom [Dataset basics: read, access, modify, write — pydicom 2.3.0 documentation](https://pydicom.github.io/pydicom/stable/tutorials/dataset_basics.html)
      1. Allows reading image
      2. Allows reading metadata
         1. Metadata will contain certain information regarding the patient
            1. Ethnicity
            2. Age
            3. Weight
            4. Anatomic region
            5. Imaging Modality
            6. More
            7. Source: [DICOM Library - Anonymize, Share, View DICOM files ONLINE](https://www.dicomlibrary.com/dicom/dicom-tags/)
      3. Allows writing metadata
         1. Can modify the existing fields
         2. Can write new fields
   2. TensorFlow
      1. Allows creation of Neural Networks
         1. Convolutional Neural Networks will be used
            1. Three convolutional layers paired with an average pooling layer
            2. Layers used to extract key features from the image
         2. Convert metadata fields into categorical fields to be used in conjunction with the
3. Performance Constraints
4. Analytical Objectives and Product Success
5. Necessary Visualization
6. Ensuring Usability
7. Interactivity and Deployment of product