**Problem statement, data and findings**

The problem is about detecting bounding boxes for lung opacity corresponding diagnosis of Pneumonia on chest radiographs (images).

The data is organized in several folders,

1. **stage\_2\_train\_images -** Contains set of raw medical images (DICOM files) for training model. The DICOM files contain a combination of header metadata as well as underlying raw image arrays for pixel data.
2. **stage\_2\_train\_labels.csv -** This CSV file contains detailed information about the labels (Patient Id, bounding boxes for lung opacity and target 1 or 0 indicate the presence of abnormality i.e. Pneumonia)
3. **stage\_2\_detailed\_class\_info.csv** - This CSV files contains information regarding three possible classes in the data, namely *normal, lung opacity* and *no lung opacity* – *not normal*
4. **stage\_2\_test\_images -** Contains set of raw medical images (DICOM files) for testing the model. The files contain a combination of header metadata as well as underlying raw image arrays for pixel data
5. **GCP Credits Request Link - RSNA.txt** -
6. **stage\_2\_sample\_submission.csv** -

All lung opacities may not attribute to Pneumonia as the Pneumonia is one of the several diseases that can occur on a chest radiograph. A radiograph may contain one or more than1 bounding boxes for any possible Pneumonia case.