Data for this challenge comes in **HDF** file format. This format was created by the HDF group (website here <https://www.hdfgroup.org/solutions/hdf5/>) and heavily leveraged in multiple industries (aerospace, automotive, government, etc)

* To view this data, you can use the HDF group’s HDF viewer application (available here <https://www.hdfgroup.org/downloads/hdfview/>)
* There are many open source libraries to access HDF data.. examples
  + Python - <https://www.h5py.org/>
  + Java - <https://wiki-bsse.ethz.ch/display/JHDF5/JHDF5+%28HDF5+for+Java%29>
  + C++ - <https://support.hdfgroup.org/HDF5/doc/cpplus_RM/index.html>
  + Rust - <https://github.com/aldanor/hdf5-rust>
  + Etc

Data recorded from the machines is documented in channels (shown in the layer below)

For each of the data files, the sampling rate recorded from the file is documented in the metadata

* NUMBER OF PARAMETERS
* NUMBER OF SCANS
* SAMPLE\_RATE

No subject matter expertise is needed to analyze this data. The goal of this competition is to use statistical methods (rather than use subject matter expertise) to detect anomalies in the data.

Utilize the ground truth XLS file to help identify data files of interest that may help focus on channels