Community Resources for RF Dataset Sharing

ISART 2022

RFDataFactory: One-stop Resource for Datasets for the Wireless Community

Kaushik Chowdhury
Professor, Department of Electrical and Computer Engineering
Associate Director for the Institute for the Wireless IoT
Northeastern University
krc@ece.neu.edu

Introduction

- Who am I? Professor at Northeastern University. Performer on DARPA RFMLS, IARPA SCISRS, NSF PI for PAWR project office, NSF Colosseum, NSF RFDataFactory
- RFDataFactory is a platform for accessing and sharing RF-centric datasets, software application programming interfaces and tutorials for collecting and processing data from experimental testbeds and simulations. Supported by the NSF Award #2120447.
- Recently concluded workshop: https://workshop.rfdatafactory.com

12

Datasets

7

APIs

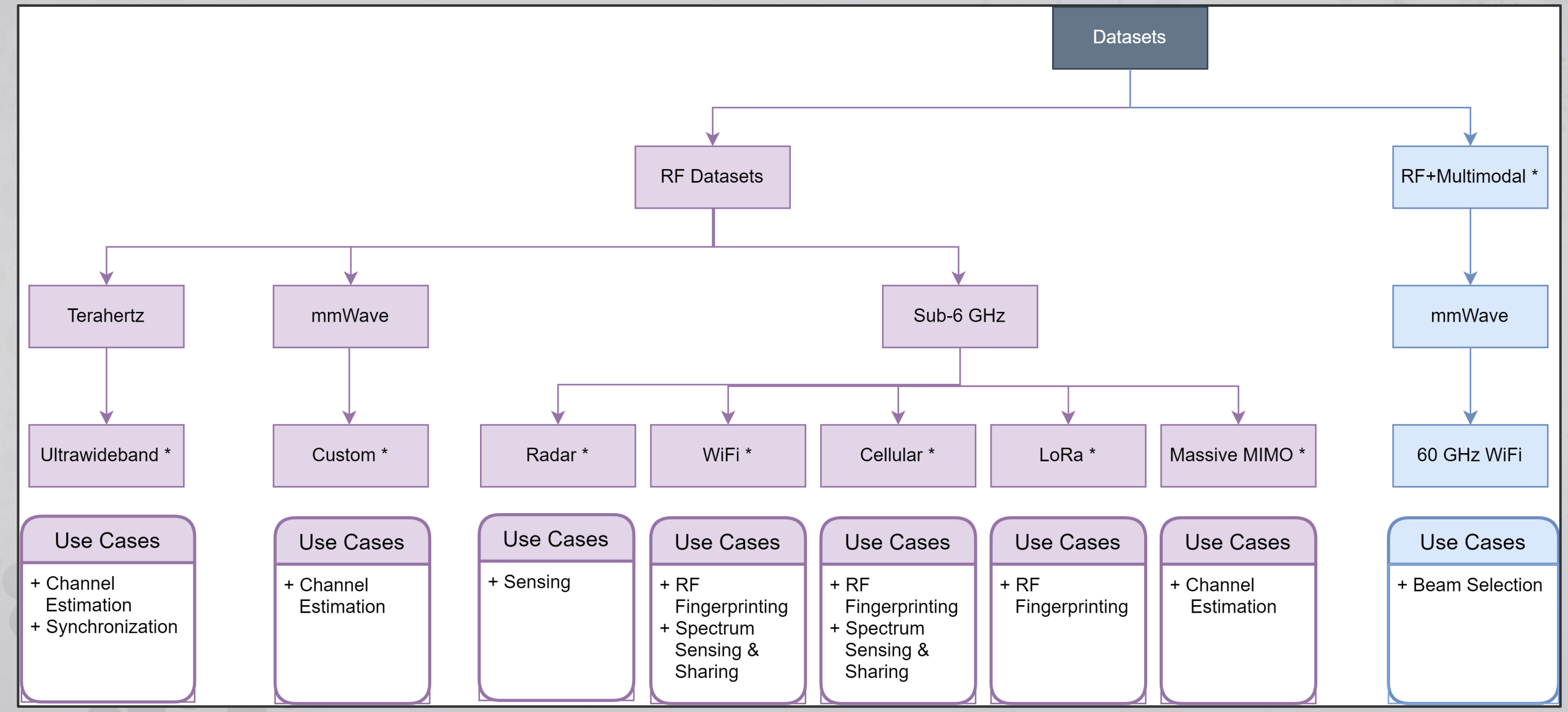
5

Tutorials

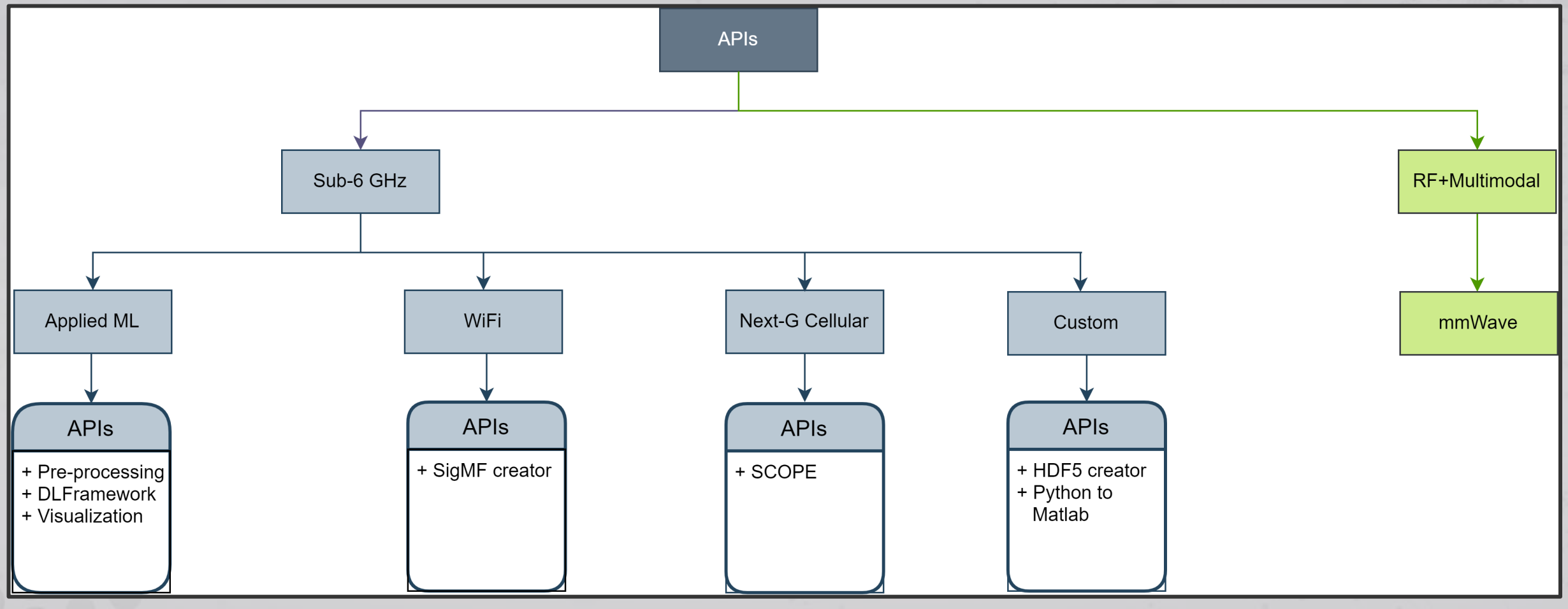
67

Access Visits

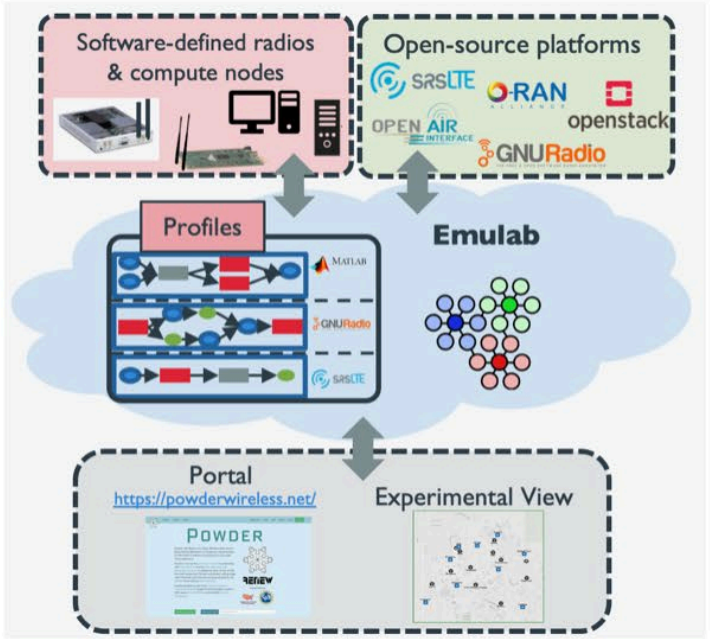
What Datasets Are Available in RFDataFactory?



What Software Resources Are Available in RFDataFactory?



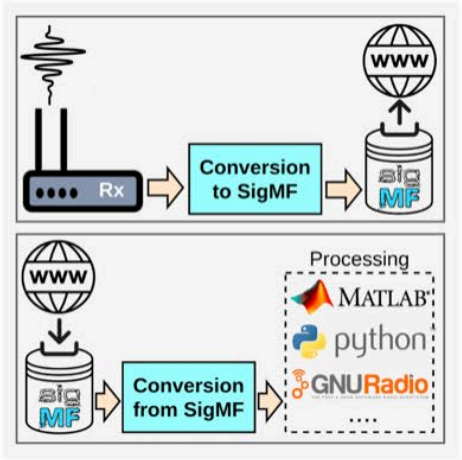
What Software Resources Are Available in RFDataFactory?



Data Collection on the POWDER PAWR Platform

Guillem Reus Muns

A step by step example on how to collect data using the POWDER platform.

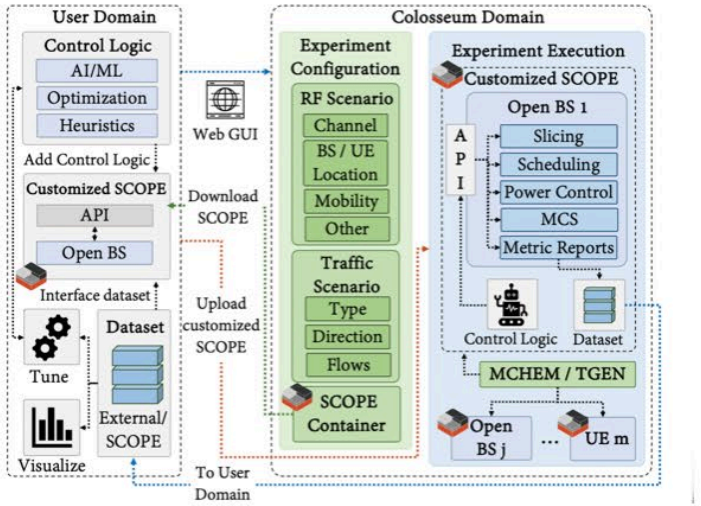


SigMF Creator

Nasim Soltani

Step by step guidline to create SigMF.

WATCH TUTORIAL



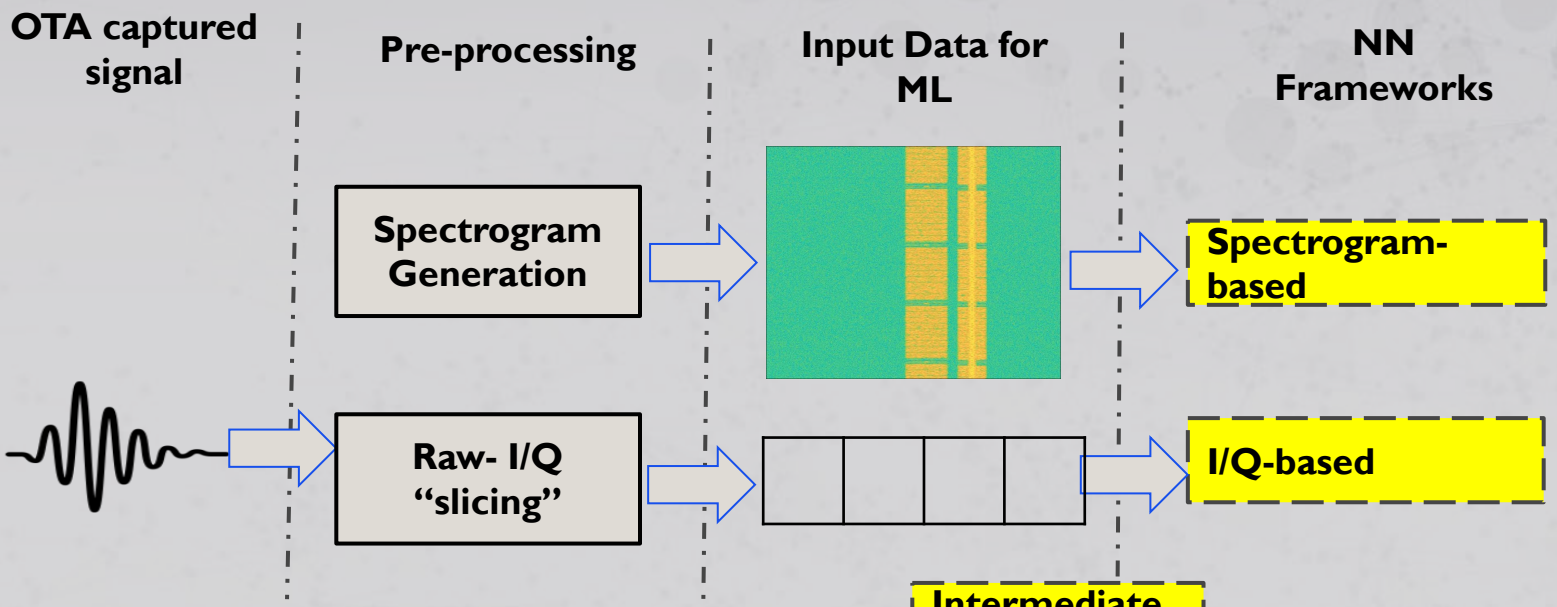
SCOPE

SCOPE is a development environment for softwarized and virtualized NextG cellular networks based on srsRAN.

WATCH TUTORIAL

Thoughts on Spectrum Data Sharing

- Transformations at the source:
  - IQ data into another usable form, such as spectrograms
- Share features not raw data
- Is SigMF ready for the future?
  - Challenges in capturing time varying and adaptive signals



Dataset Description:

We are releasing two datasets ***a) Dataset #1 : recordings of raw IQ samples collected from over-the-air transmissions of 16 USRP X310 transmitter radios ; b) Dataset #2: recordings of demodulated IQ symbols collected after equalizing over-the-cable transmissions of 16 IQ imbalance configurations.*** In both the datasets, each recording consists of two files: a metadata file and a dataset file. The dataset file is a binary file of digital samples, and the metadata file contains information that describes the dataset. Our metadata and data format is an extension of, and compatible with the SigMF specifications .

- **Dataset #1 :** It consists of recordings of collected raw IQ samples from 16, high-end X310 USRP SDRs with the same B210 radio as a receiver. The recordings are categorized into different folders with folder name "xxft", where xx represents the transmitter-receiver separation distance in feet. Each recording has a dataset file with an extenstion of *'.sigmf-data'*, and a metadata file with an extension of *'.sigmf-meta'*. These files are named in a specific format for more intuitive understanding. For example, the dataset file "***WiFi_air_X310_3123D7B_2ft_run1***" represents
  - ***WiFi :*** --> IEEE802.11a standard-compliant WLAN frame
  - ***air :***--> medium of transmission
  - ***X310 :***--> the type of USRP radio
  - ***3123D7B :*** --> device serial ID
  - ***2ft:*** --> the transmitter-receiver separation distance in feet
  - ***run1 :*** --> the recording number
  - ***sigmf-data/sigmf-meta :*** --> the extension of dataset file/metadata file