**Instructions for Accessing Datasets**

**Main Datasets**:

1. [CoAID: COVID-19 Healthcare Misinformation Dataset](https://ui.adsabs.harvard.edu/abs/2020arXiv200600885C/abstract)
   1. Go to <https://github.com/cuilimeng/CoAID> to access Github repository containing datasets
   2. Click green “Code” button and select “Download ZIP” to download zip file
2. [FakeHealth](https://github.com/EnyanDai/FakeHealth)
   1. Go to <https://github.com/EnyanDai/FakeHealth> to access Github repository containing datasets
   2. Click green “Code” button and select “Download ZIP” to download zip file
3. [COVID-19 Fake News](https://github.com/diptamath/covid_fake_news/tree/main/data)
   1. Go to <https://github.com/diptamath/covid_fake_news/tree/main/data> to access Github repository containing datasets
   2. Click “[covid\_fake\_news](https://github.com/diptamath/covid_fake_news)” link to navigate to parent repository
   3. Click green “Code” button and select “Download ZIP” to download zip file
   4. Delete all folders except “data”
   5. All dataset files will be found in the “data” folder
4. [COVID-19 Tweets Dataset](https://github.com/lopezbec/COVID19_Tweets_Dataset)
   1. Go to <https://github.com/lopezbec/COVID19_Tweets_Dataset> to access Github repository containing datasets
   2. Click green “Code” button and select “Download ZIP” to download zip file

**Supplementary Datasets:**

1. [Covid-19 Vaccine Tweets](https://www.kaggle.com/kaushiksuresh147/covidvaccine-tweets)
   1. Go to <https://www.kaggle.com/kaushiksuresh147/covidvaccine-tweets>
   2. Click “Download (77 MB)” button to download dataset files
2. [COVID Fake News Dataset](https://zenodo.org/record/4282522#.YKXcGZNKhhE)
   1. Go to <https://zenodo.org/record/4282522#.YKXiYJNKhhF>
   2. Click “Download” button under files section to download dataset file
3. [COVID-19 Tweets Dataset](https://github.com/lopezbec/COVID19_Tweets_Dataset)
   1. Go to <https://zenodo.org/record/3901741#.YKXirZNKhhE>
   2. Click “Download” button(s) under files section to download dataset file(s)

**Note:**

Many of the datasets contain tweet IDs and not the content of the tweet itself. In order to convert tweet IDs back into the content of the tweets in a CSV or JSON format, you will need to use a tool called [Hydrator.](https://github.com/DocNow/hydrator)

Download the latest release of Hydrator from [here](https://github.com/DocNow/hydrator/releases). Select file format which is compatible with your operating system.

1. .dmg for macOS
2. .exe for Windows
3. .deb for Linux

Before using Hydrator, you will need to have a Twitter account. In order to convert tweet IDs within Hydrator, copy only the tweet IDs from your dataset file into a Google Spreadsheet or any other Spreadsheet software and download the spreadsheet in .csv or .txt file format. This .csv or .txt file can then be interpreted and converted by Hydrator. Below are helpful resources to learn how to use Hydrator. Even though resources may use older versions of Hydrater, instructions/steps will still remain the same.

1. <https://covid.dh.miami.edu/2020/06/11/hydrating-tweetsets/>
2. <https://www.youtube.com/watch?v=6QMRIEMczyY&ab_channel=EdSummersEdSummers>
3. Read hydrating section: <https://programminghistorian.org/en/lessons/beginners-guide-to-twitter-data#hydrating>