Sergiu Buhatel, Research Assistant – Computing/ML

Graph - Internetworking

Table of Contents

[1 Introduction 1](#_Toc206233542)

[2 Data 1](#_Toc206233543)

[2.1 Raw Data 1](#_Toc206233544)

[2.2 Preprocess Data 2](#_Toc206233545)

[3 Frontend 2](#_Toc206233546)

[4 GitHub 2](#_Toc206233547)

[5 References 2](#_Toc206233548)

# Introduction

# Data

## Raw Data

The raw data is stored as CSV files for each year from 2017 to 2023.

For example, in **2017** the files are:

* 2017\_industries.csv
* 2017\_Network.csv
* Top50\_2017.csv
* Top100\_2017.csv
* Top200\_2017.csv

Similarly, in **2018** the files are:

* 2018\_industries.csv
* 2018\_Network.csv
* Top50\_2018.csv
* Top100\_2018.csv
* Top200\_2018.csv

In general, the data follows the structure:

<year>/<year>\_industries.csv

<year>/<year>\_Network.csv

<year>/Top50\_<year>.csv

<year>/Top100\_<year>.csv

<year>/Top200\_<year>.csv

The only exception is **2023**, where Top50\_2023.csv and Top100\_2023.csv are missing, and the available file is named with a lowercase prefix: top200\_2023.csv.

## Preprocess Data

Open a command prompt and launch:

**create\_all\_data.bat**

This will create all data based on raw data from **.\frontend\public\data\<year>**

It makes the assumption that all directories of .\frontend\public\data are the years and each directory includes raw data in it.

# Frontend

# GitHub

The code is available on GitHub at <https://github.com/sergiubuhatel/ra2/tree/main/graph>

# References