Scientific Research Visualization: A Case Study on Covid-19 Literature

Gianluca Tasciotti 16/03/2021

Table of Contents

1. Introduction

2. Text mining

3. Clustering

4. Visualization platform

5. Conclusions and future works

Scientific Research Visualization: A Case Study on Covid-19 Literature (G. Tasciotti — 2021)

Mission

Tracking the evolution of the scientific research along the time using a visualization platform

Scientific Research Visualization: A Case Study on Covid-19 Literature (G. Tasciotti — 2021)

1. Introduction

Papers per year in CORD-19
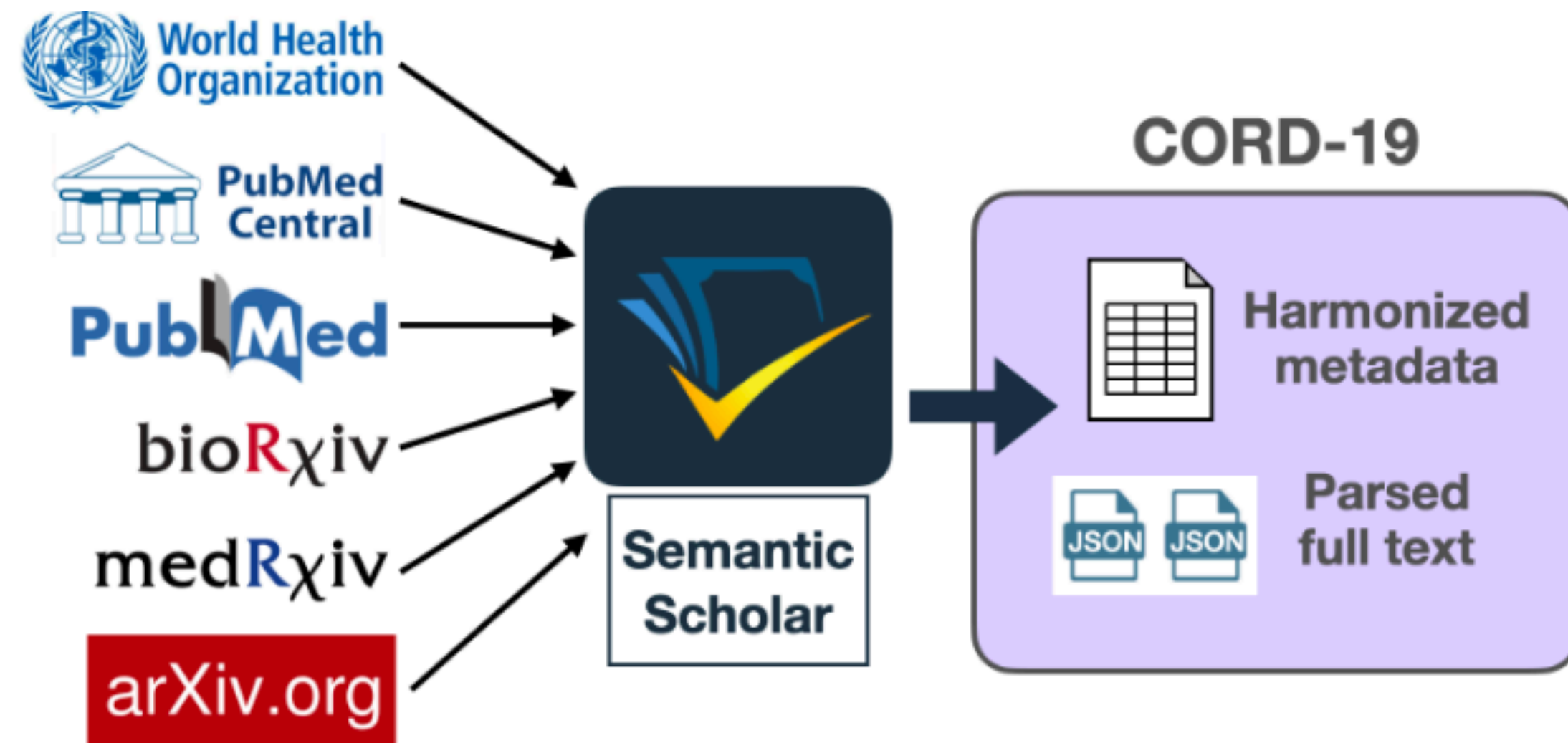
Papers and preprints are collected from different sources through Semantic Scholar. Released as part of CORD-19 are the harmonized and deduplicated metadata and full text JSON. [1]
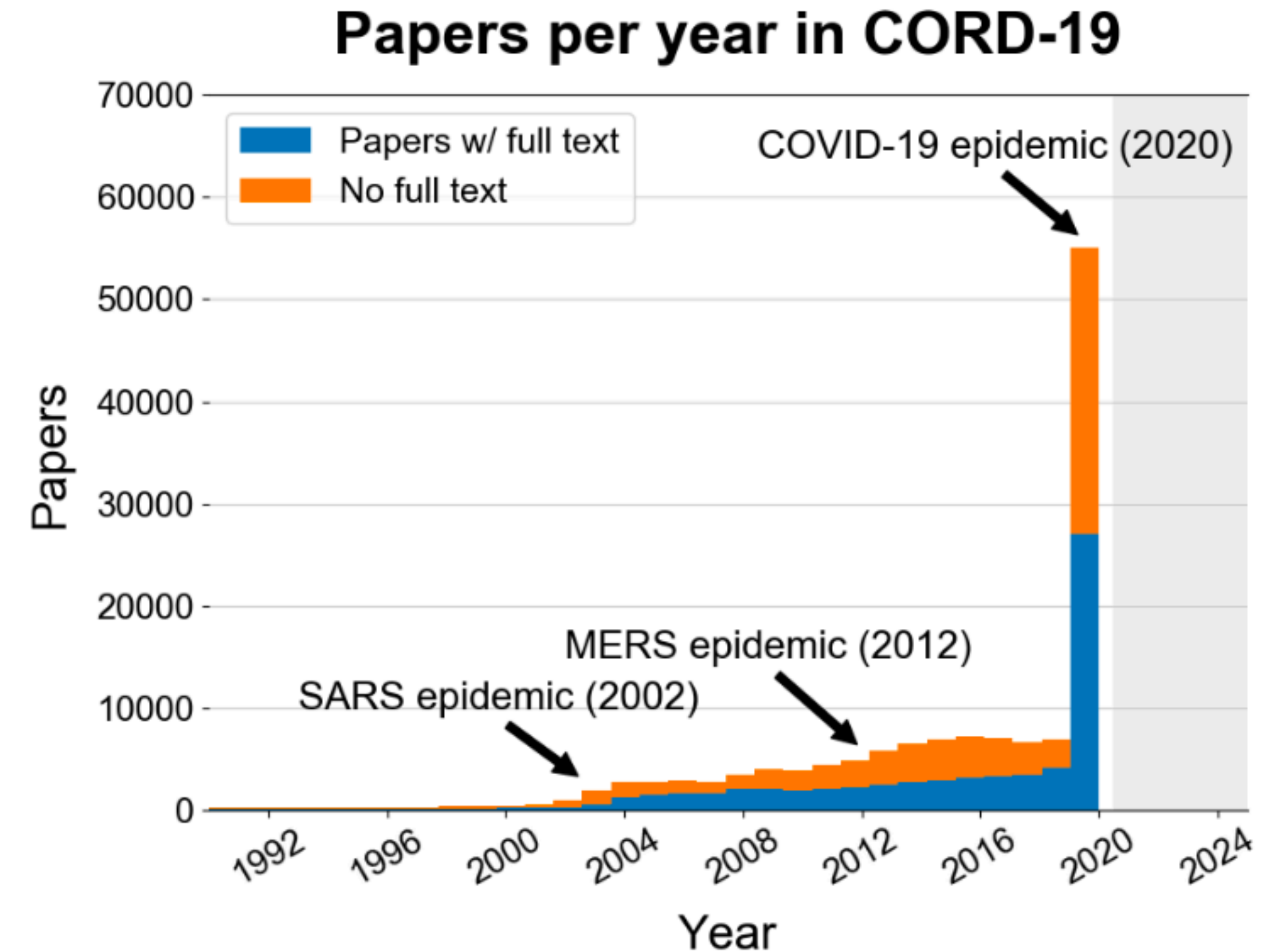
The distribution of papers per year in CORD-19. A spike in publications occurs in 2020 in response to COVID-19. [1]

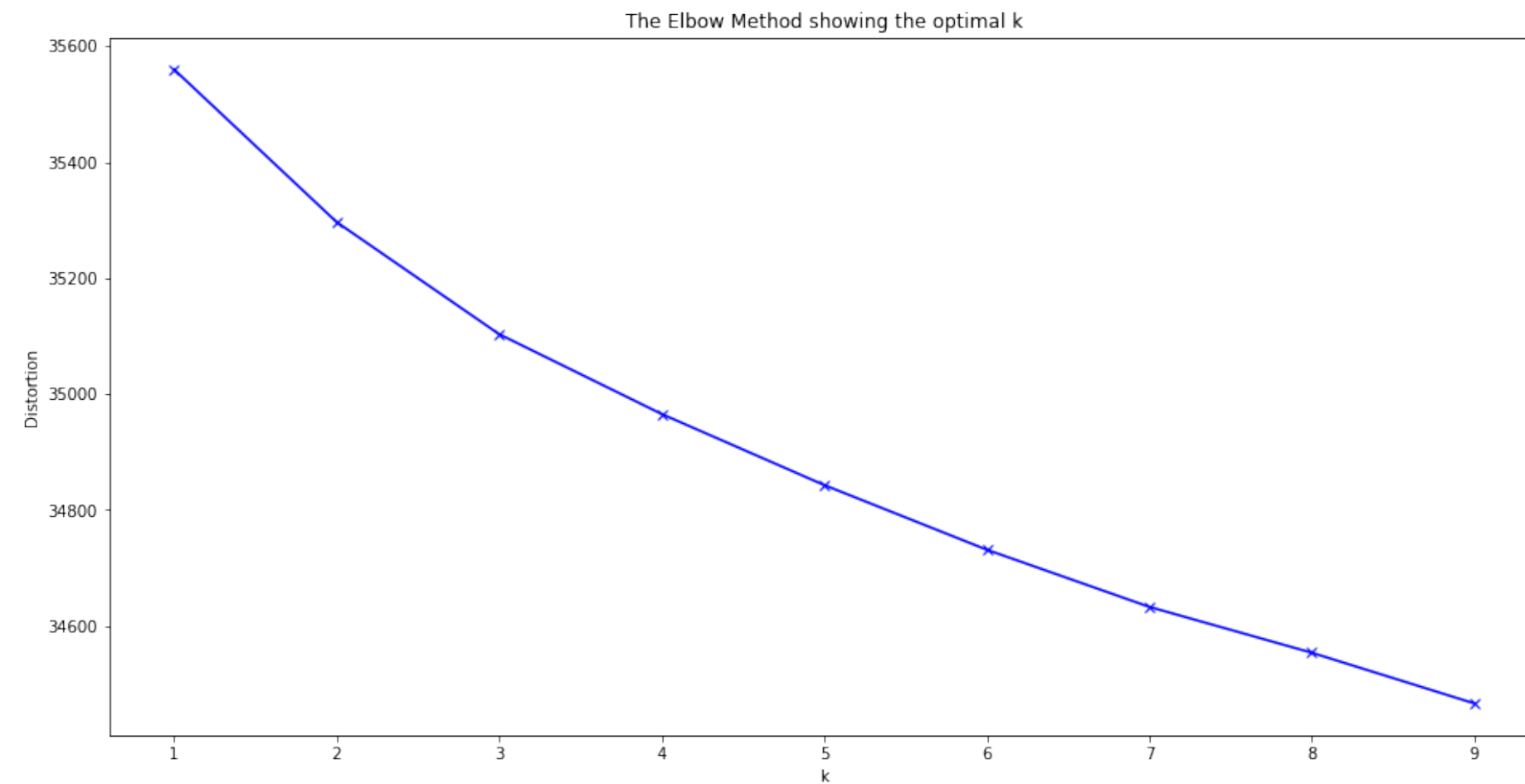**Scientific Research Visualization: A Case Study on Covid-19 Literature (G. Tasciotti — 2021)**

[1]   Paper: https://www.aclweb.org/anthology/2020.nlpcovid19-acl.1/.

2. Dataset
CORD-19: COVID-19 Open Research Dataset [2]

- Preprocessing
  - ‣ Tokenizing abstracts
  - ‣ Customized stopwords
  - ‣ Lemmatization
  - ‣ TF-IDF

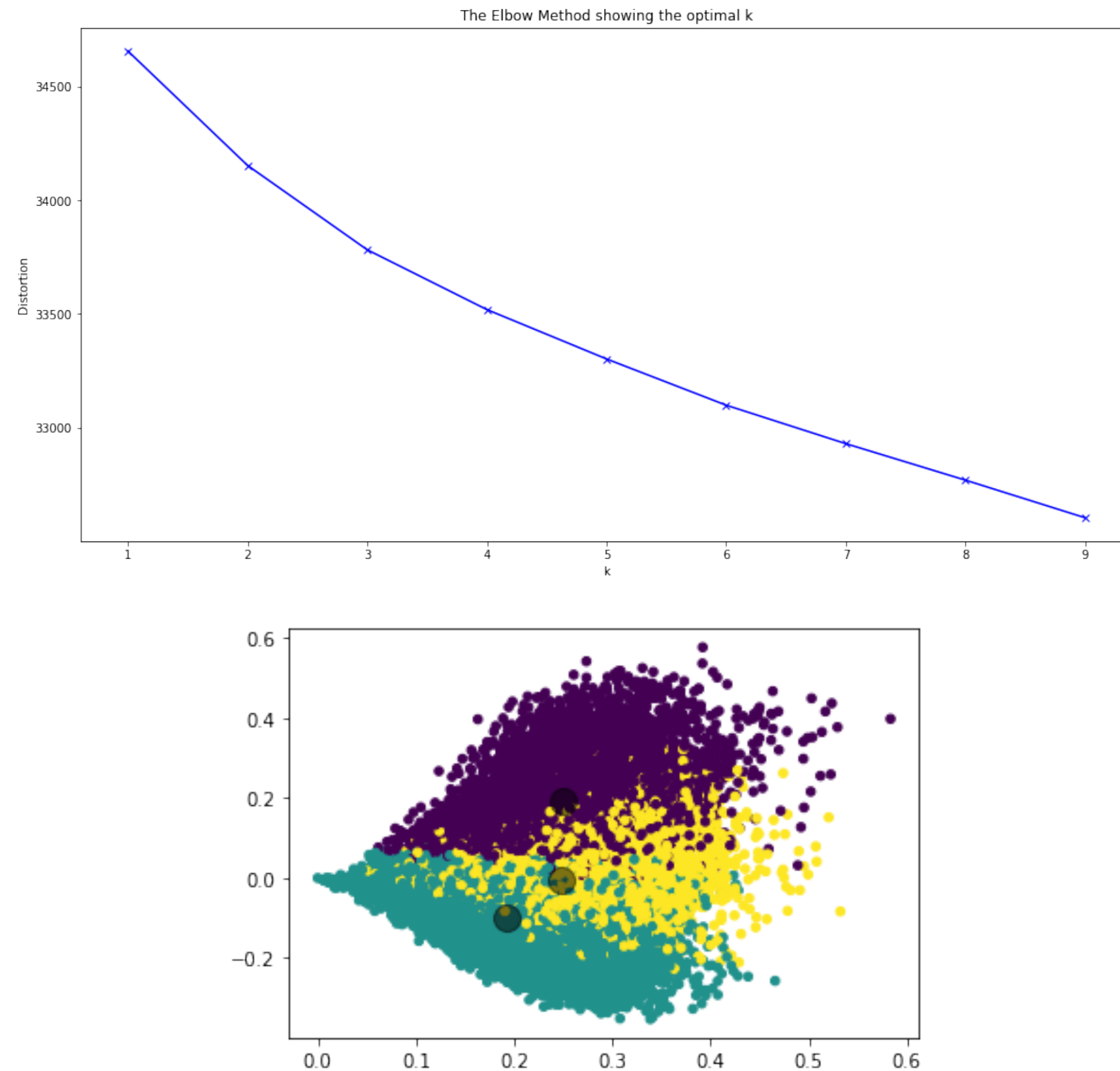**Scientific Research Visualization: A Case Study on Covid-19 Literature (G. Tasciotti — 2021)**

[2]   Dataset: *https://ai2-semanticscholar-cord-19.s3-us-west-2.amazonaws.com*
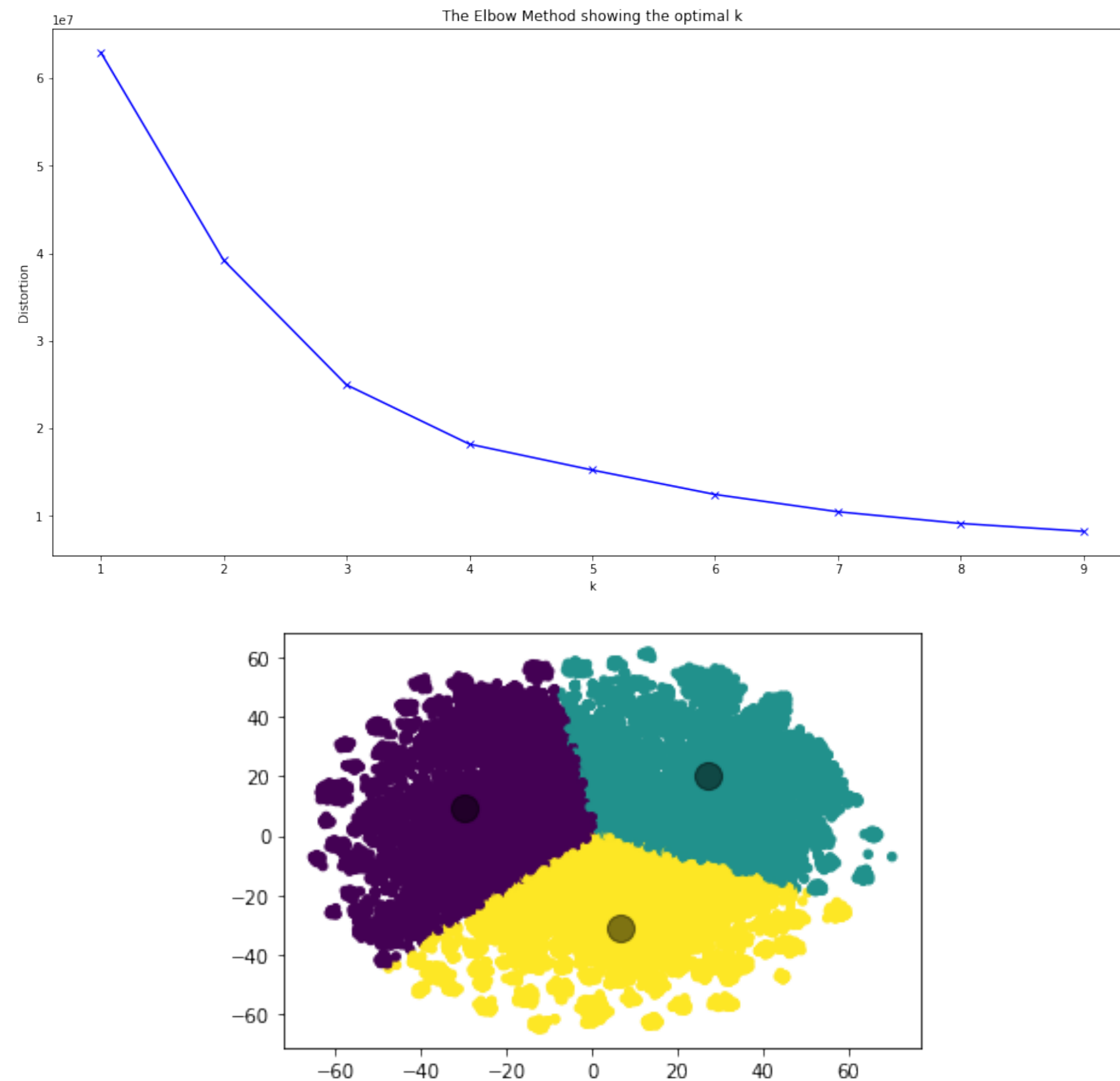
3. Clustering - Original
CORD-19: COVID-19 Open Research Dataset

Scientific Research Visualization: A Case Study on Covid-19 Literature (G. Tasciotti — 2021)
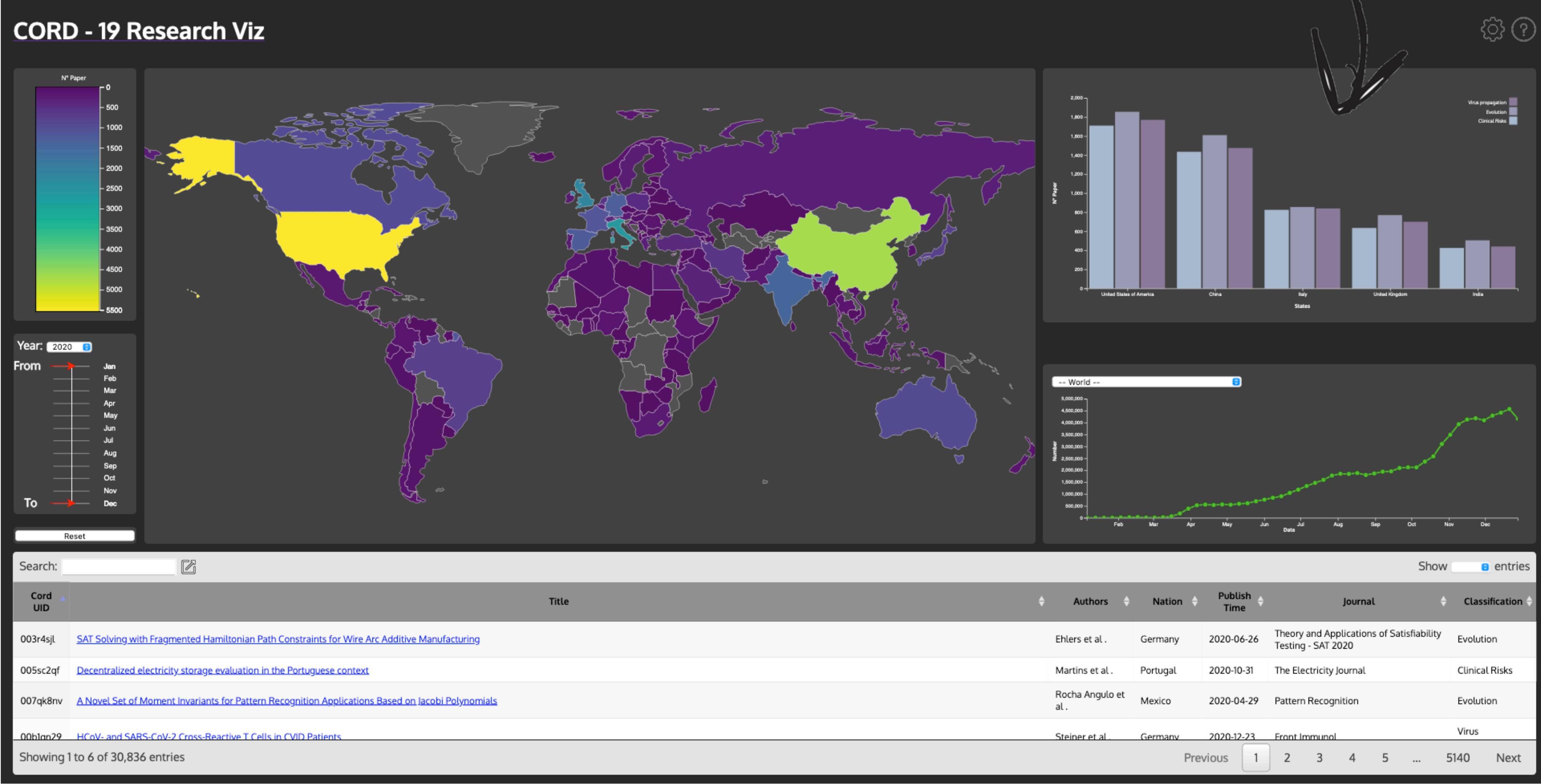
3. Clustering - SVD
CORD-19: COVID-19 Open Research Dataset

Scientific Research Visualization: A Case Study on Covid-19 Literature (G. Tasciotti — 2021)

3. Clustering - t-SNE
CORD-19: COVID-19 Open Research Dataset

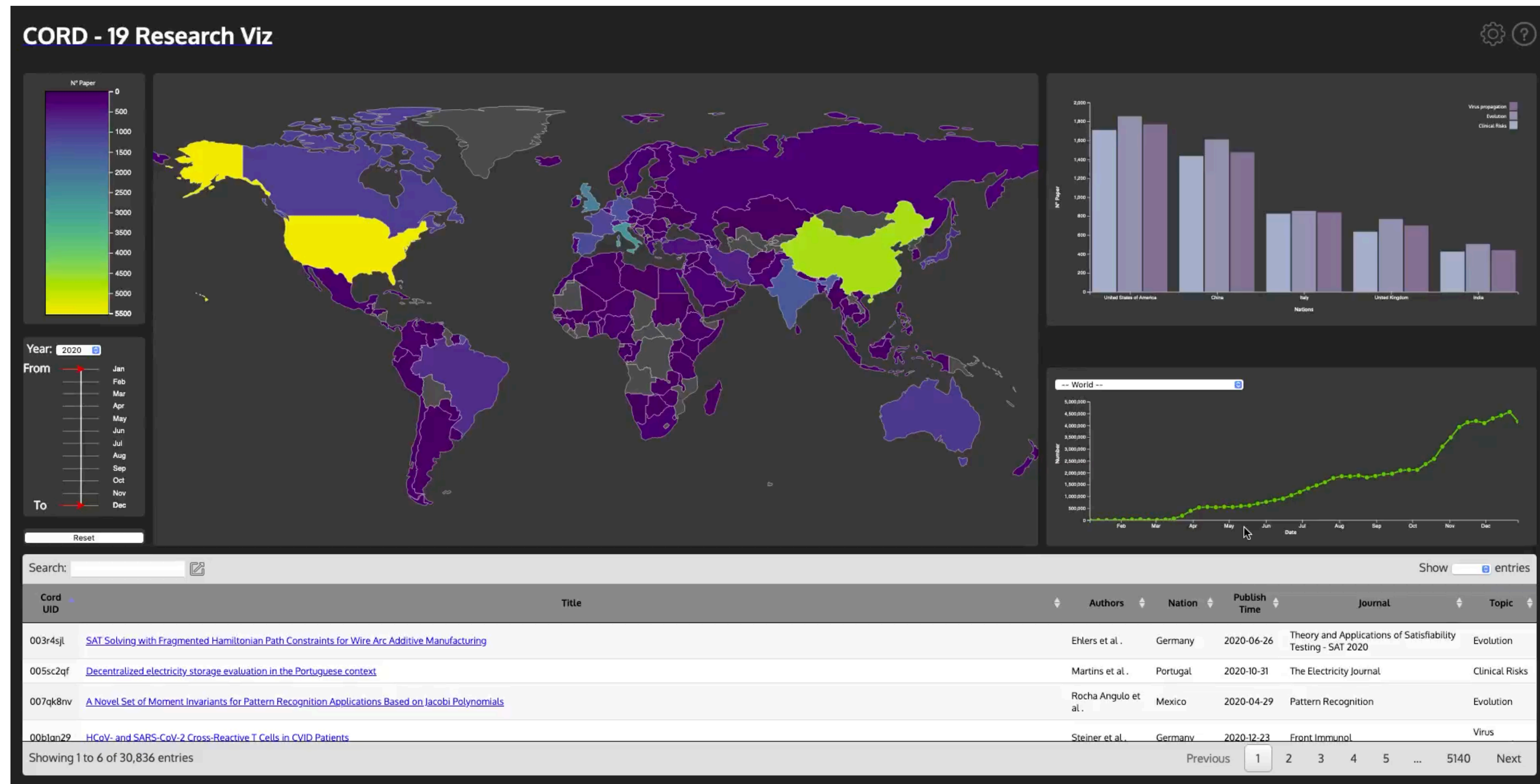**Scientific Research Visualization: A Case Study on Covid-19 Literature (G. Tasciotti — 2021)**
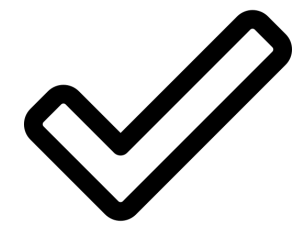
4. Visualization platform

Scientific Research Visualization: A Case Study on Covid-19 Literature (G. Tasciotti — 2021)
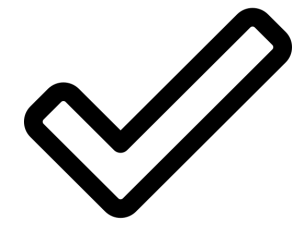
4. Visualization platform - Demo

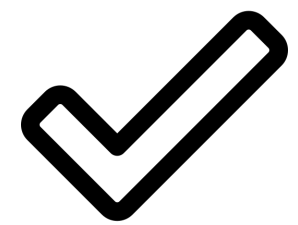**Scientific Research Visualization: A Case Study on Covid-19 Literature (G. Tasciotti — 2021)**
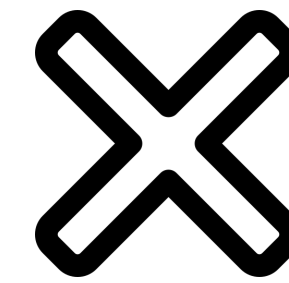
5. Discussion and conclusion

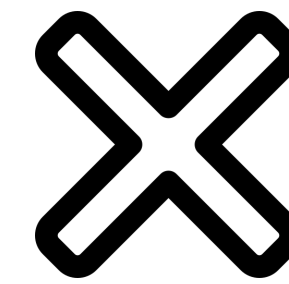✓ A different point of view to see the scientific research
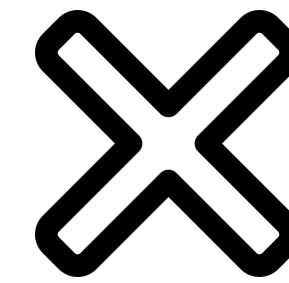
✓ Topic classification

✓ Possible generalization

✗ Huge amount of data

✗ Serverless

✗ Applying a NN

Scientific Research Visualization: A Case Study on Covid-19 Literature (G. Tasciotti — 2021)

Thanks!