Visualize GitHub Social Network

Asep Trisna Setiawan

23221111
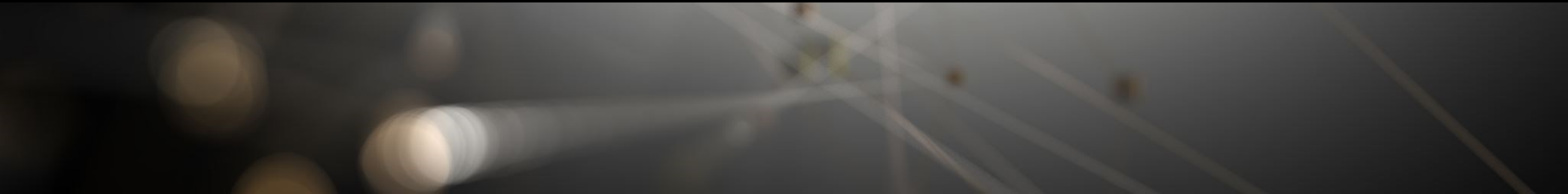
EXPLORER

proses_data.py 3 ✕   .env   test.py

proses_data.py > ...

∨ DEGREE CENTRALITY DISTRIBUTION
  > 📁 dataset
  🎚 .env
  {..} author.json
  📊 github_edges_1.csv
  📊 github_target_1.csv
  🐍 proses_data.py   3
  🐍 test.py

```python
 1  import pandas as pd
 2
 3  edges = pd.read_csv("github_edges_1.csv")
 4  edges.head(10)
 5  edges = edges.astype(str)
 6  nodes = pd.read_csv("github_target_1.csv")
 7  nodes.head(10)
 8  # print(nodes)
 9  nodes.id = nodes.id.astype(str)
10  nodes.ml_target = nodes.ml_target.astype(str)
11  percent_web, percent_ml = nodes.ml_target.value_counts(normalize=True)
12  import plotly.express as px
13
14  px.histogram(data_frame=nodes, x="ml_target", histnorm="probability density")
15  edges = edges.merge(nodes, how="left", left_on="id_1", right_on="id").merge(
16      nodes, how="left", left_on="id_2", right_on="id", suffixes=("1", "2")
17  )[["id_1", "id_2", "name1", "name2", "ml_target1", "ml_target2"]]
18  edges.head(10)
19  # print(edges)
20  cross_edge = edges.query("ml_target1 != ml_target2")
```

PROBLEMS 27   TERMINAL

> TERMINAL   > POWERSHELL

```
Traceback (most recent call last):
  File "D:\23221111\Data S2\Semester 2\Jaringan Informasi dan Sosial\Degree Centrality Distribution\proses_data.py",
line 51, in <module>
    import cudf
ModuleNotFoundError: No module named 'cudf'
PS D:\23221111\Data S2\Semester 2\Jaringan Informasi dan Sosial\Degree Centrality Distribution>
```
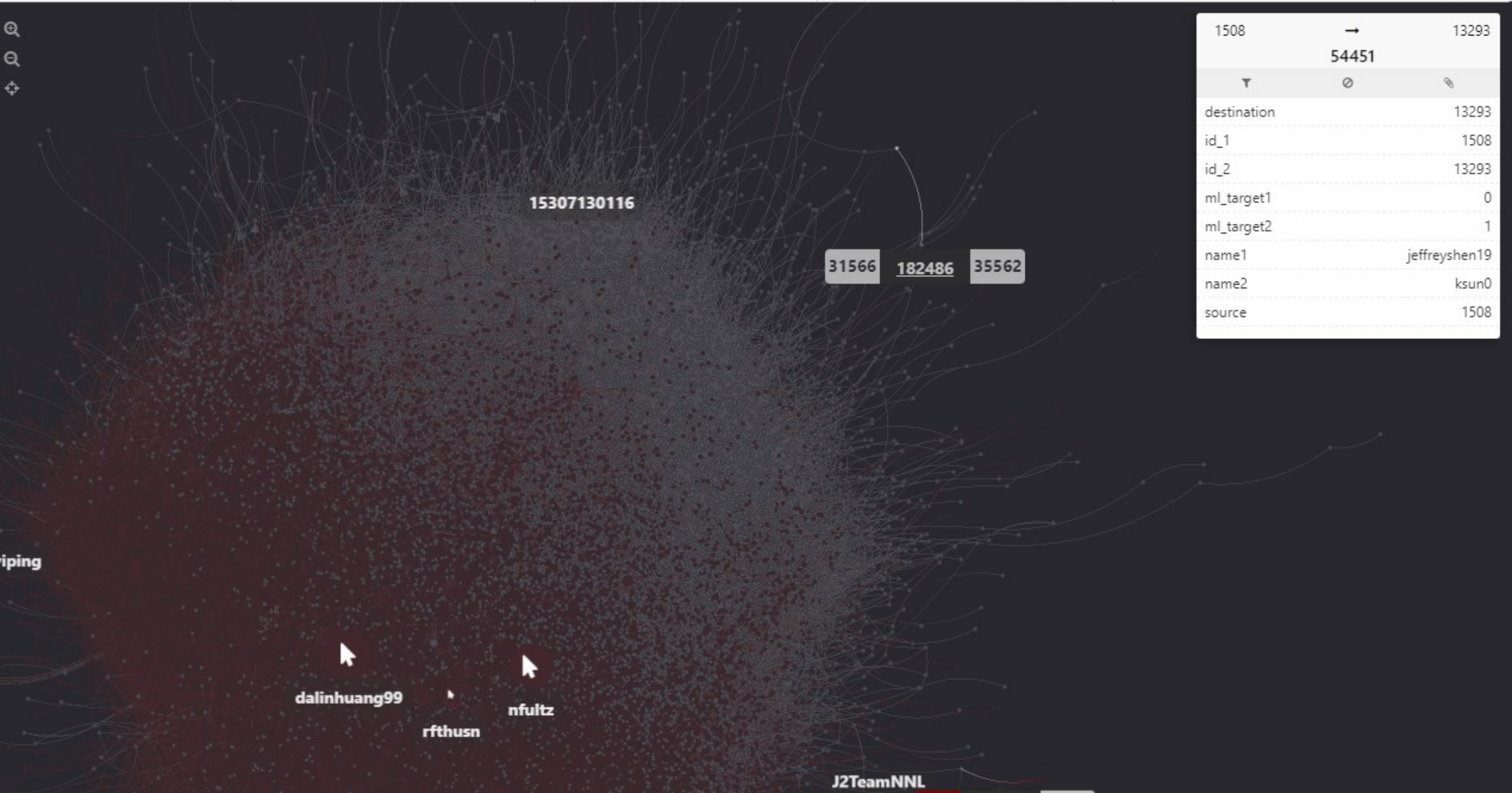
> OUTLINE

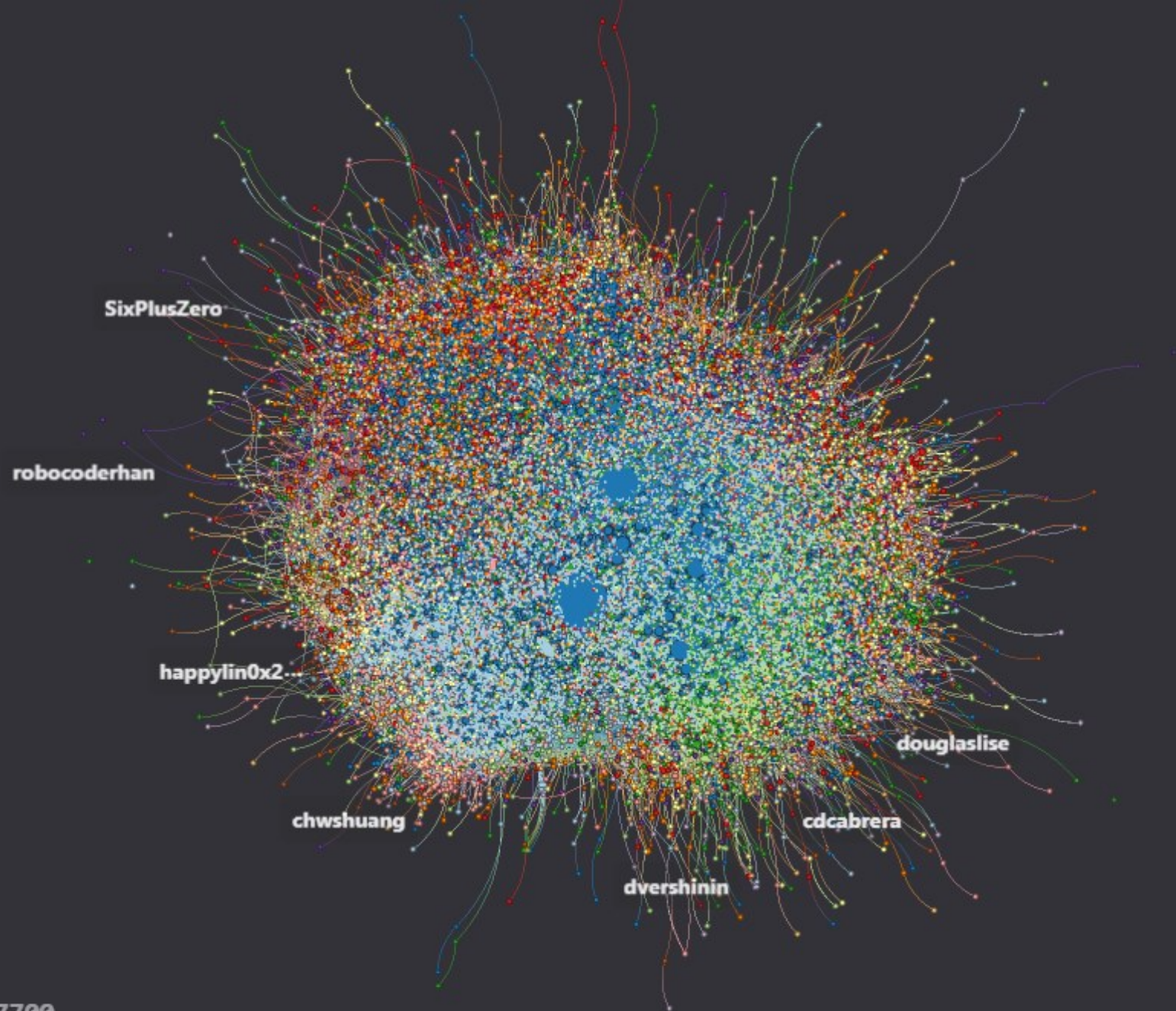⊗ 0 ⚠ 3 ⓘ 24   Ln 60, Col 20   Spaces: 4   UTF-8   CRLF   {} Python   kite: not installed   ⚠ 20 Spell   ⊘ Prettier

1508	→	13293
	54451	

▼	⊘	📎
destination		13293
id_1		1508
id_2		13293
ml_target1		0
ml_target2		1
name1		jeffreyshen19
name2		ksun0
source		1508

15307130116

31566   182486   35562

viping

dalinhuang99

rfthusn

nfultz

J2TeamNNL

SixPlusZero

robocoderhan

happylin0x2...

douglaslise

chwshuang

cdcabrera

dvershinin

NODES 37700

EDGES 289003

HIDE STANDALONE NODES

graphistry FREE TIER