

CSSS 2012

Multi-dimensional Analysis of Media

Davide Beraldo

Juan Galán Páez

Robin Lamarche-Perrin

The Data

DATABASE	Sources	Articles	Content	Time	Space	Topics
Pulse Web	200	20,000	ALL	8 years	France	1
Factiva	35,000	360,000,000	ALL	35 years	ALL	ALL
Geomedia	1 (50)	14,400 (800,000)	Abstracts	1 year	ALL	ALL
NYTimes API	1	40,000 (300,000)	Abstracts	30 years	USA	6 (ALL)

Temporal analysis

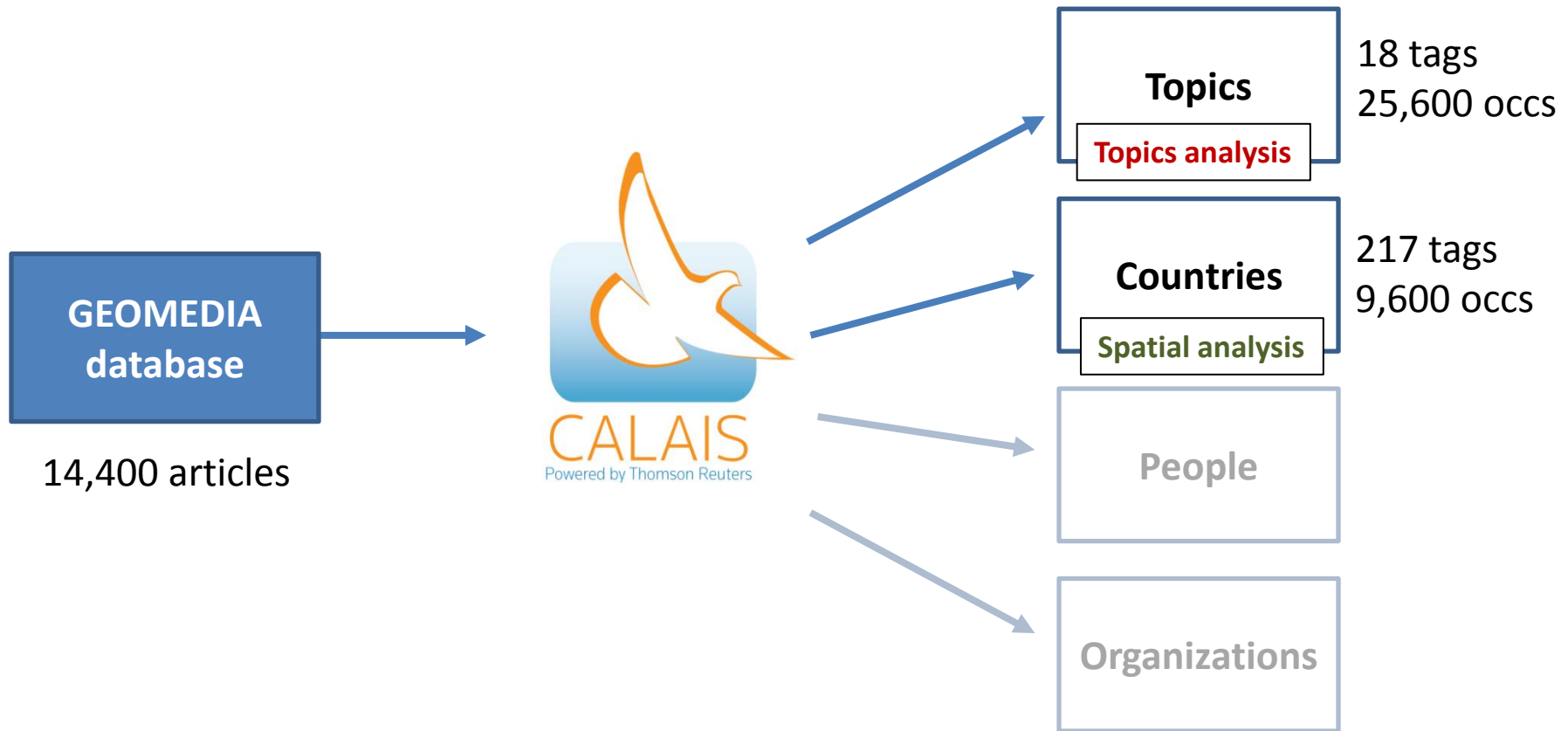
Spatial analysis

Topics analysis

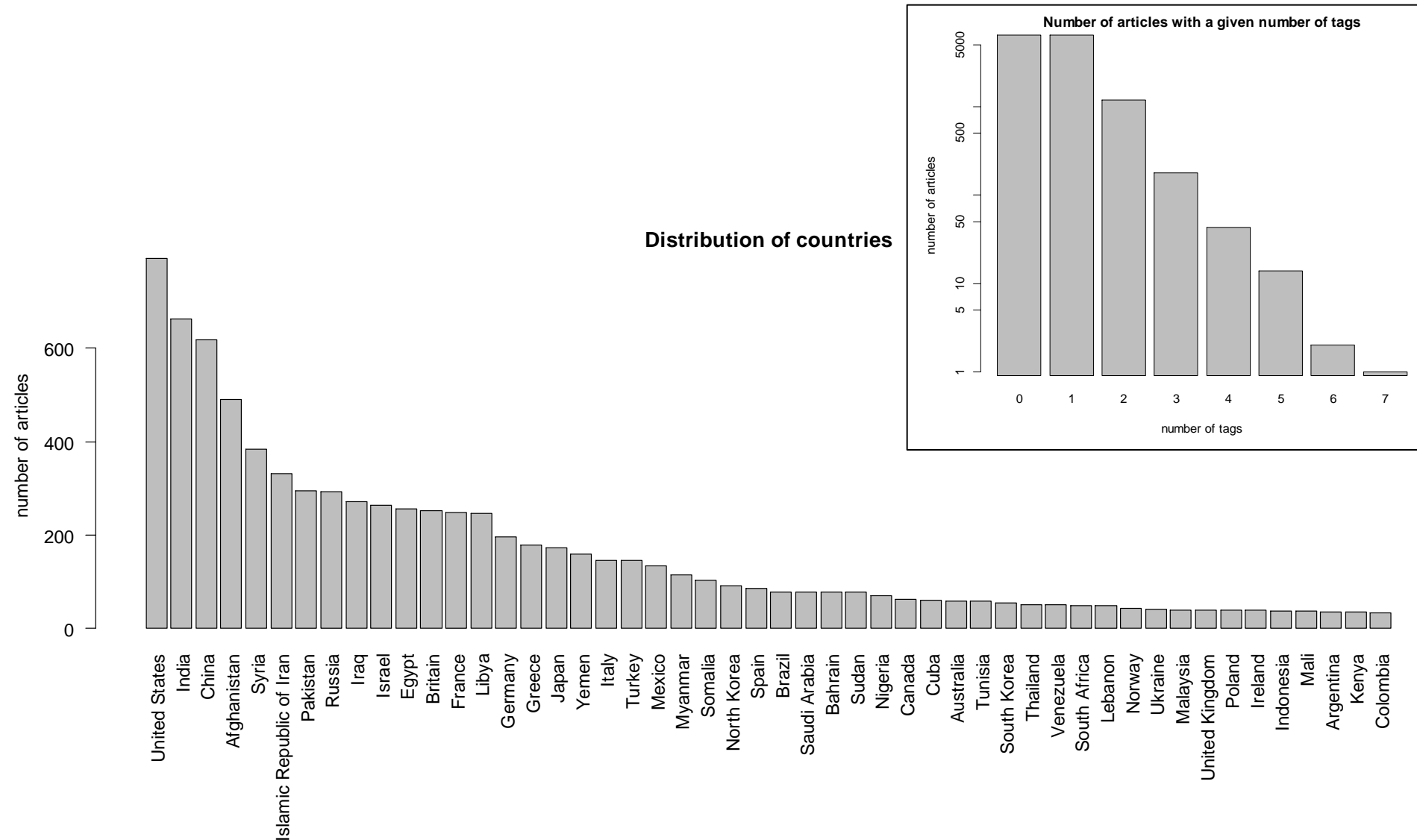
EXPERIMENT 1

Spatial Topics Analysis

Textual Analysis



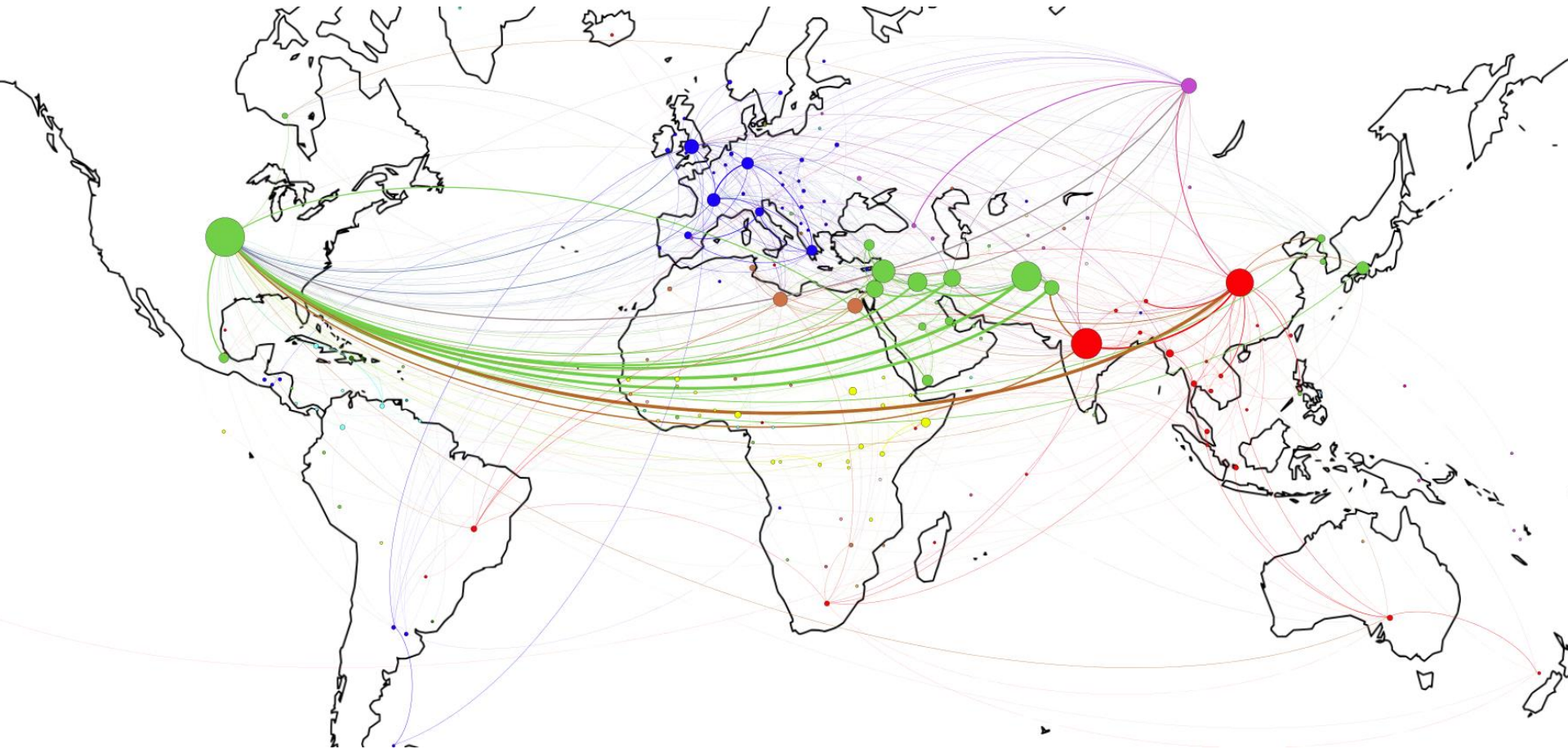
Countries from OpenCalais



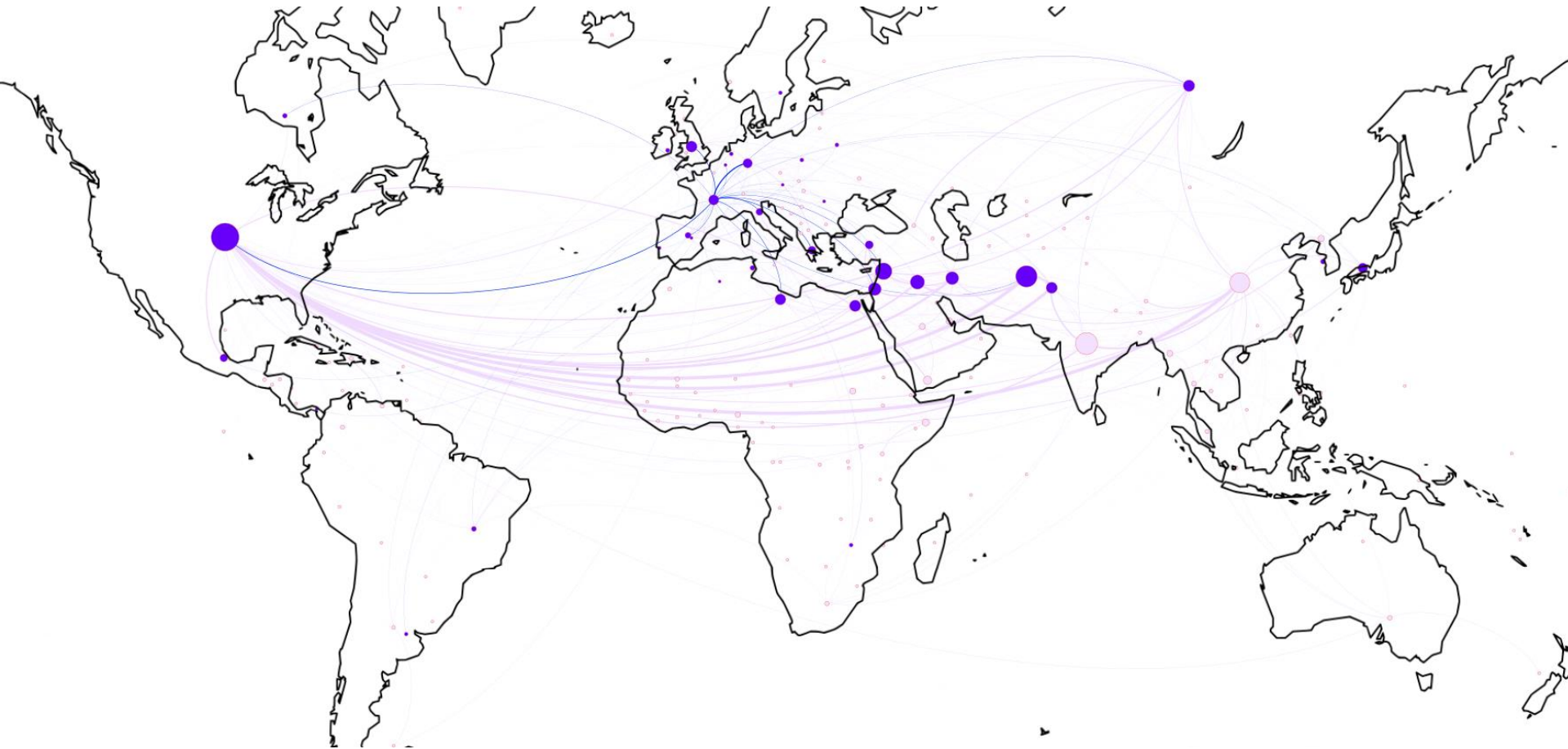
COUNTRIES RELATIONS LIST – All Topics

CHINA – USA	60
AFGHANISTAN- USA	59
PAKISTAN – USA	53
IRAN – USA	43
AFGHANISTAN – PAKISTAN	34
IRAQ – USA	34
AFGHANISTAN – PAKISTAN	32
CHINA – INDIA	29
INDIA – PAKISTAN	28
IRAN – ISRAEL	27
INDIA – USA	25
RUSSIA – USA	23
SYRIA – TURKEY	23
MEXICO – USA	23
IRAN – SYRIA	21

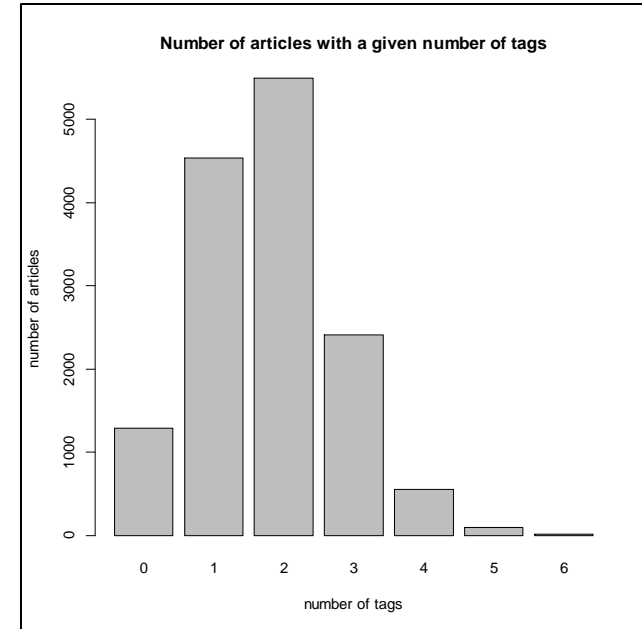
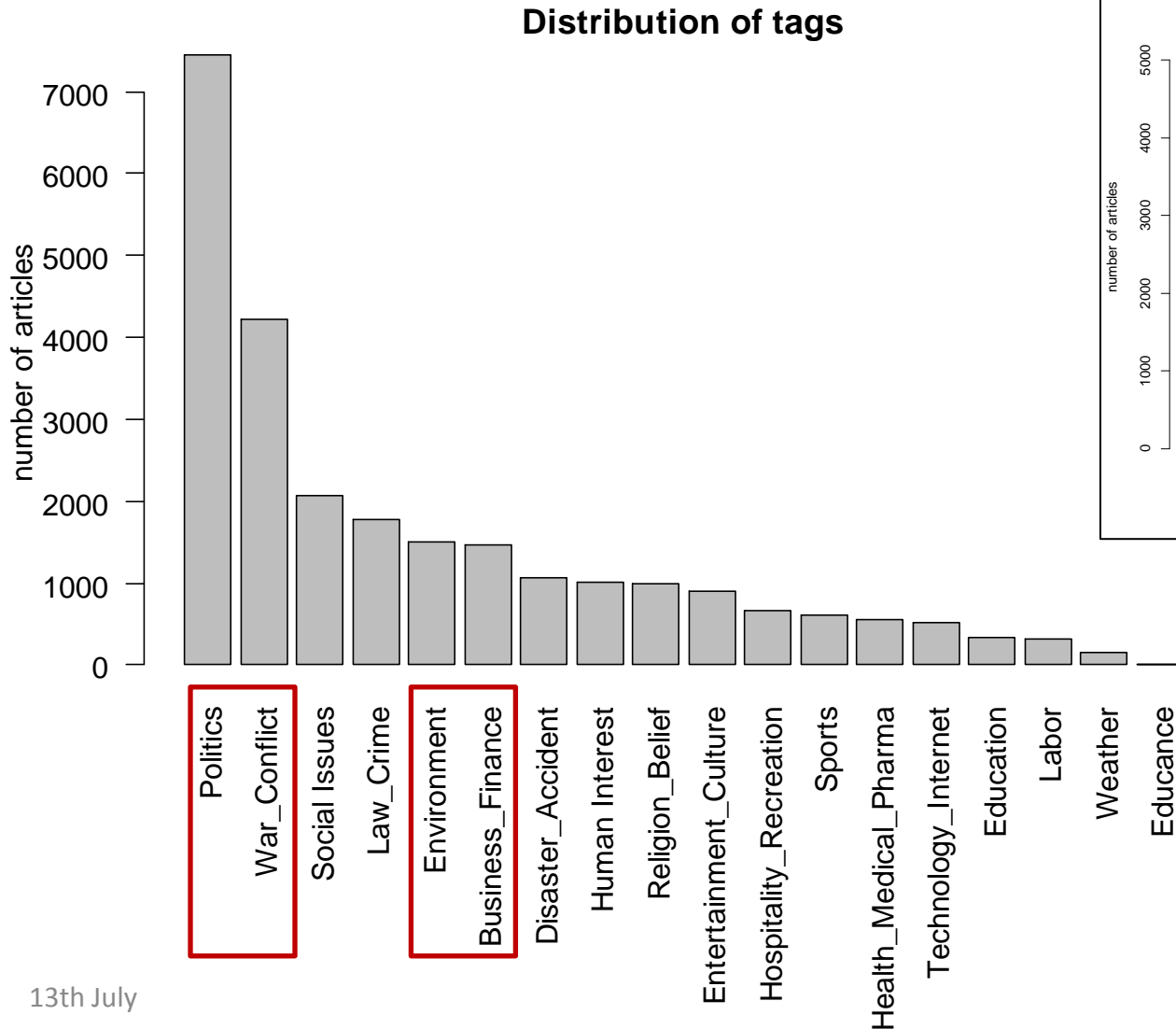
COUNTRIES RELATIONS MAP - All Topics



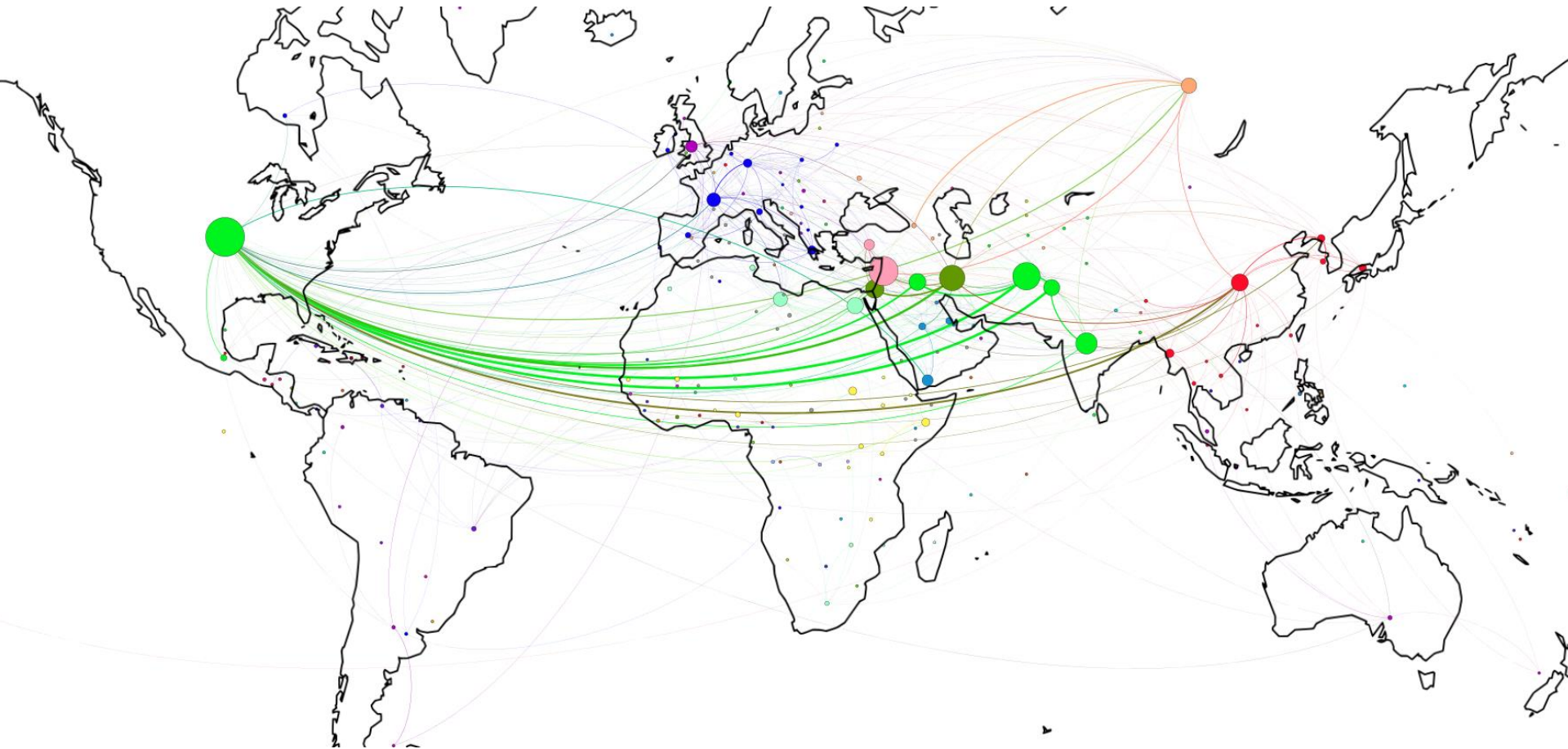
COUNTRIES RELATIONS – France



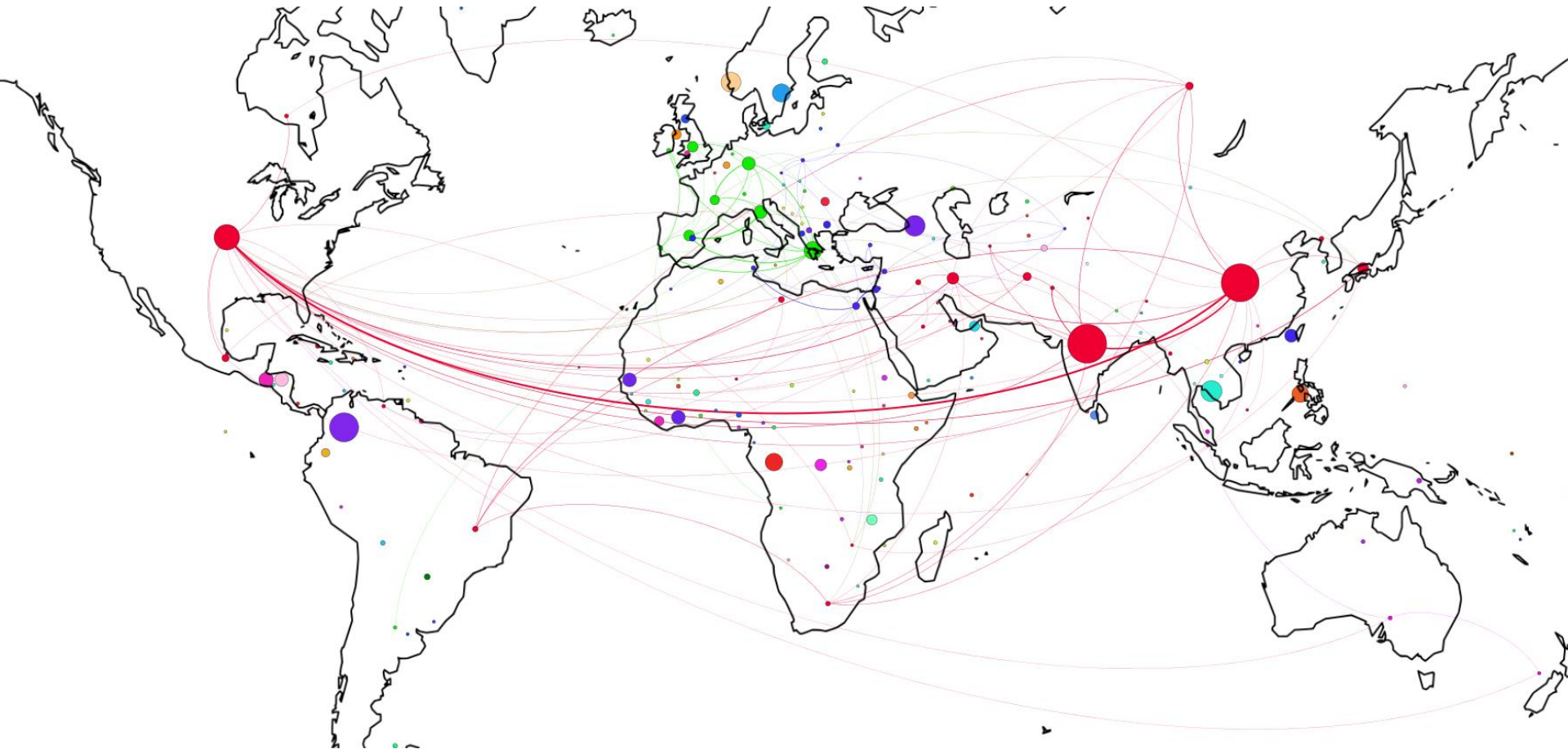
Topics from OpenCalais



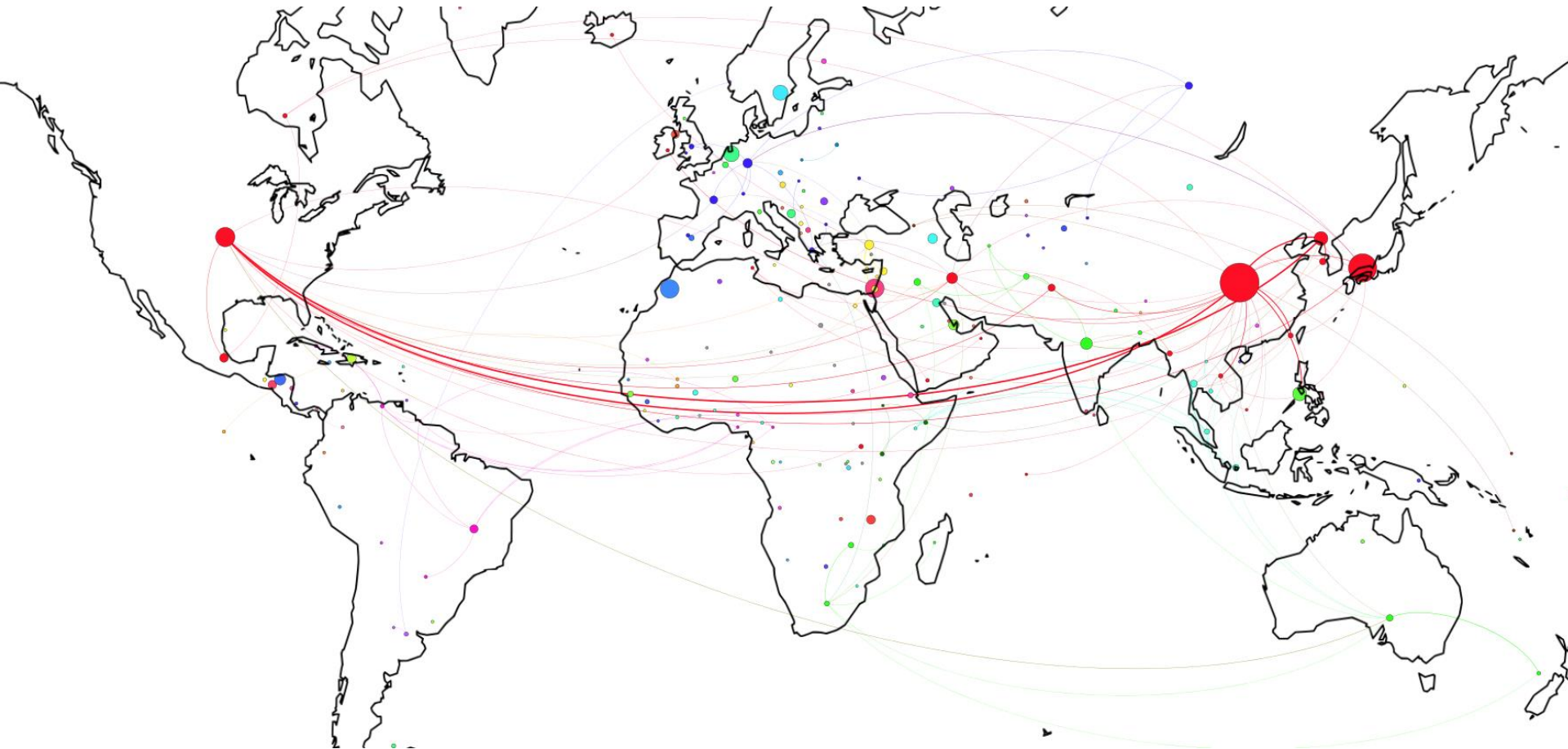
COUNTRIES RELATIONS - Politics



COUNTRIES RELATIONS - Business



COUNTRIES RELATIONS - Environment



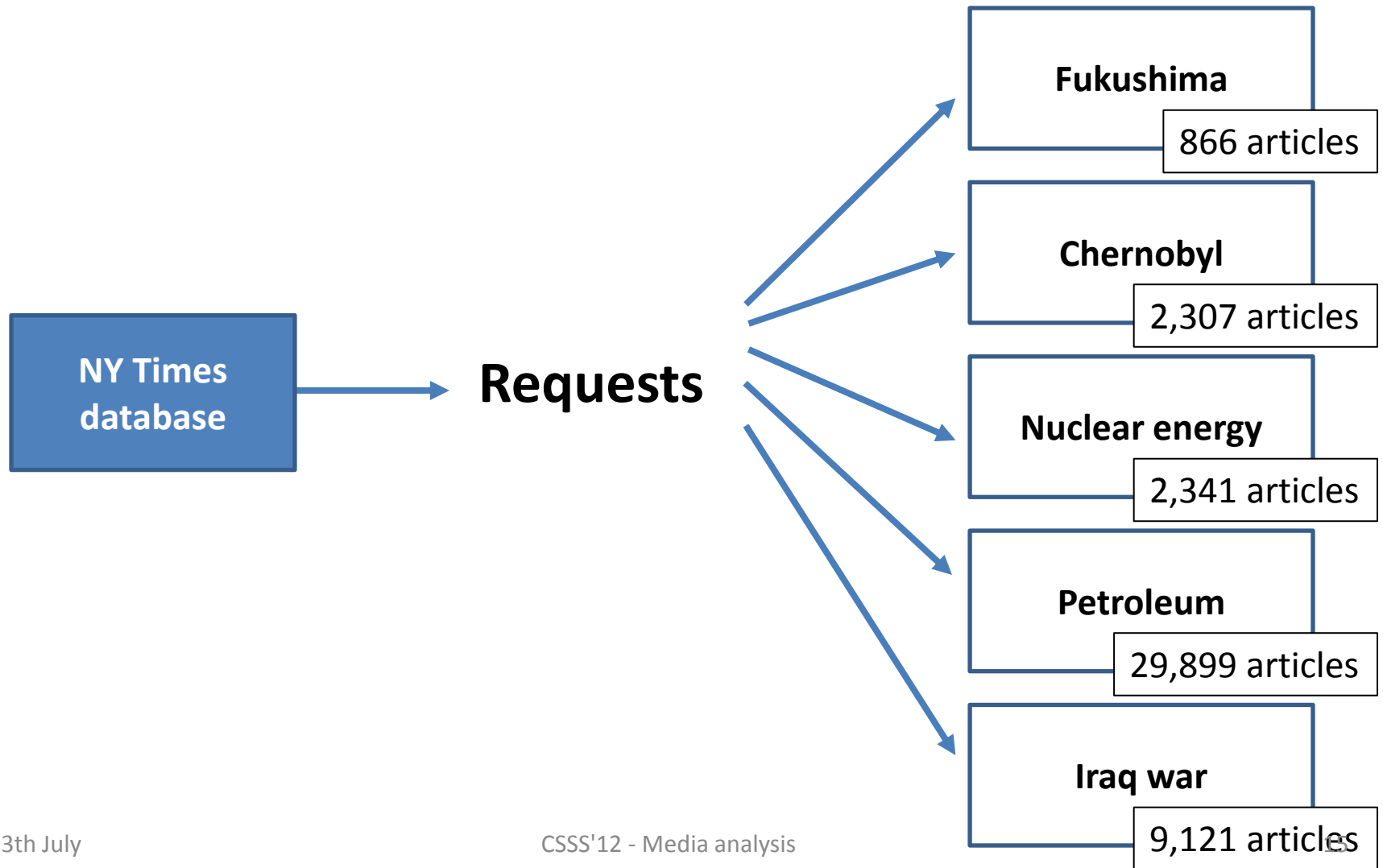
COUNTRIES RELATIONS - War



EXPERIMENT 2

Temporal Topics Analysis

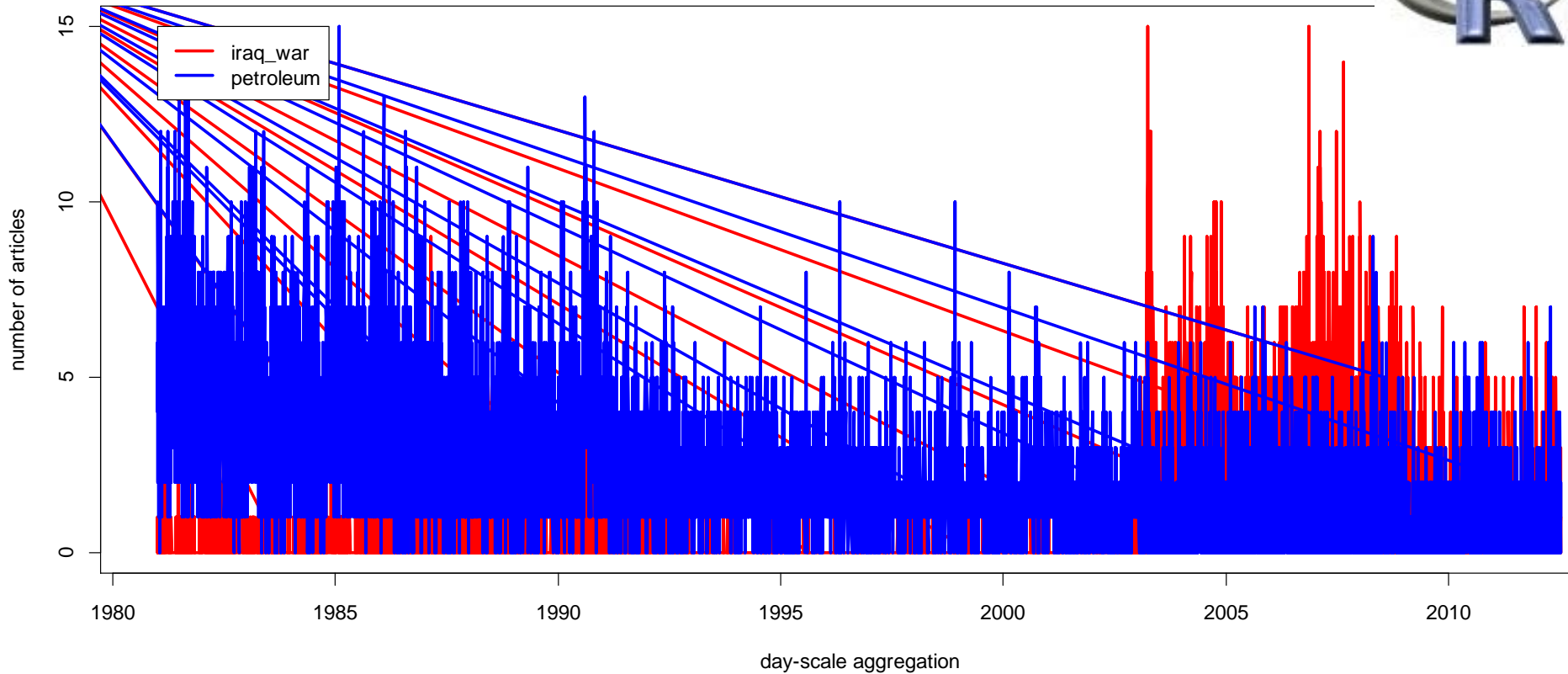
Data Retrieval

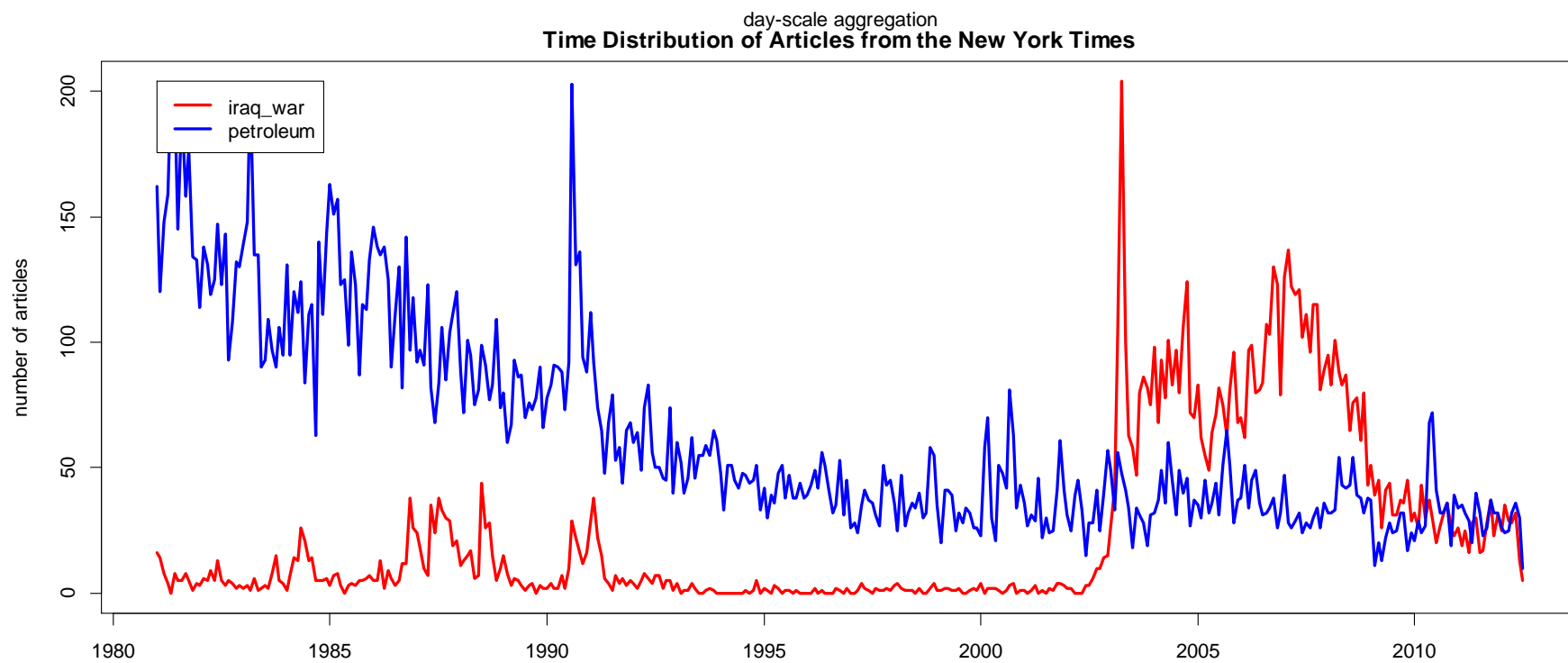
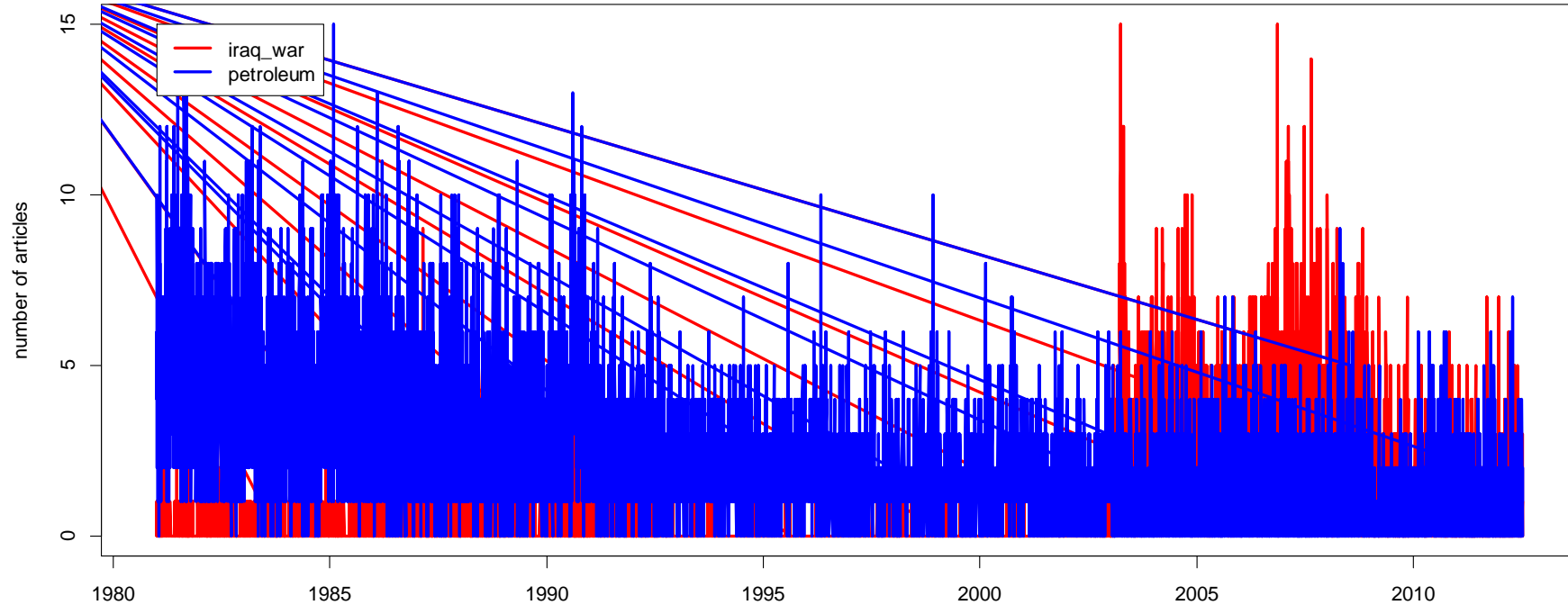


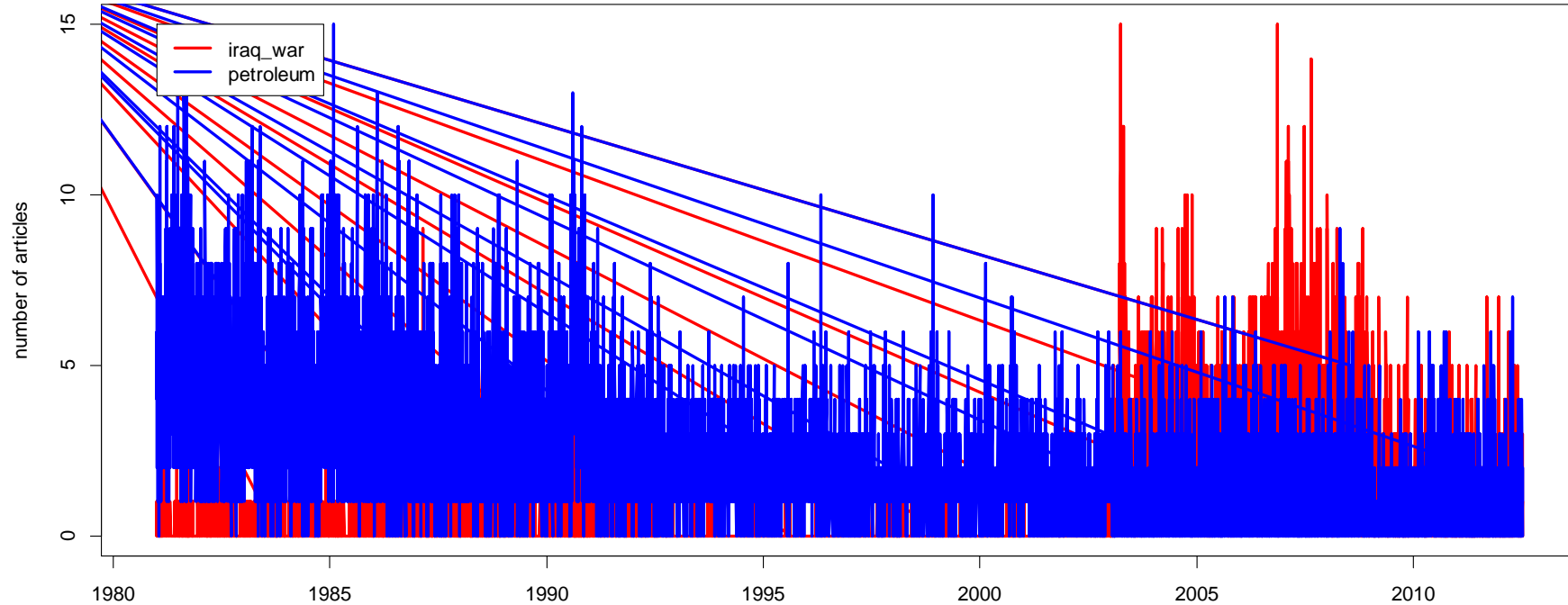
Time distribution



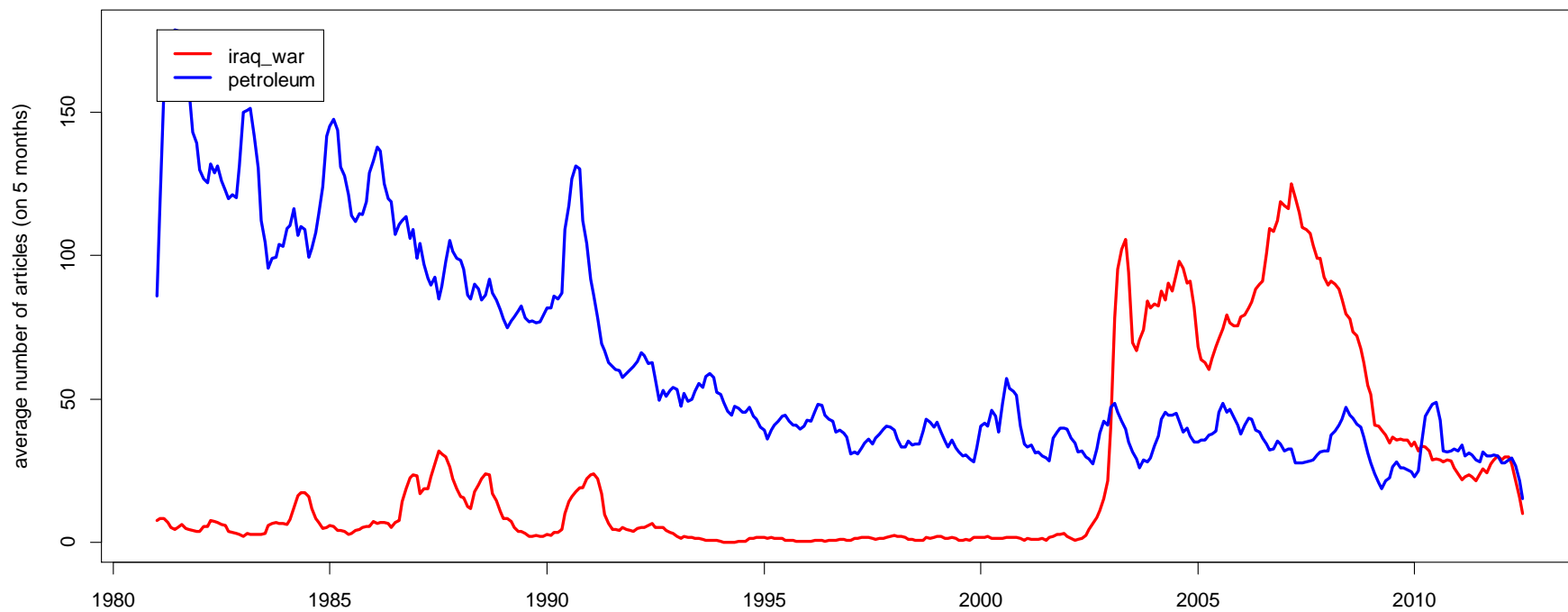
Time Distribution of Articles from the New York Times

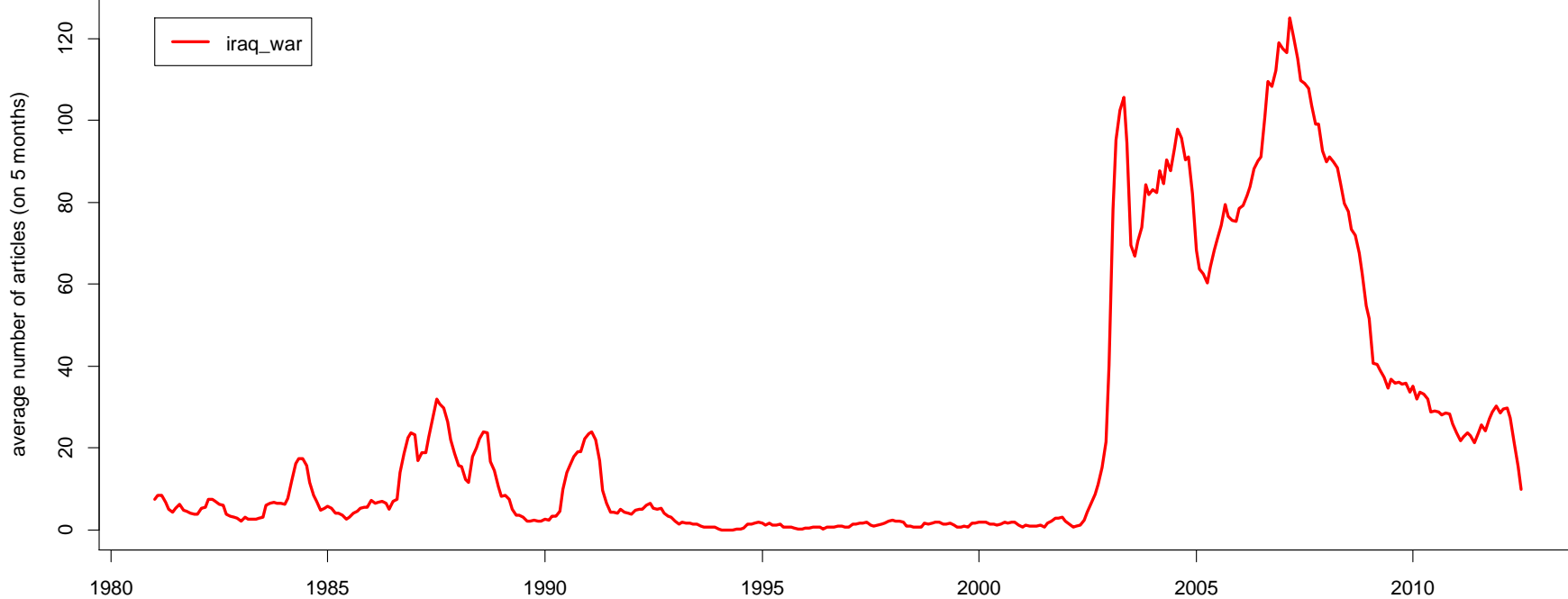




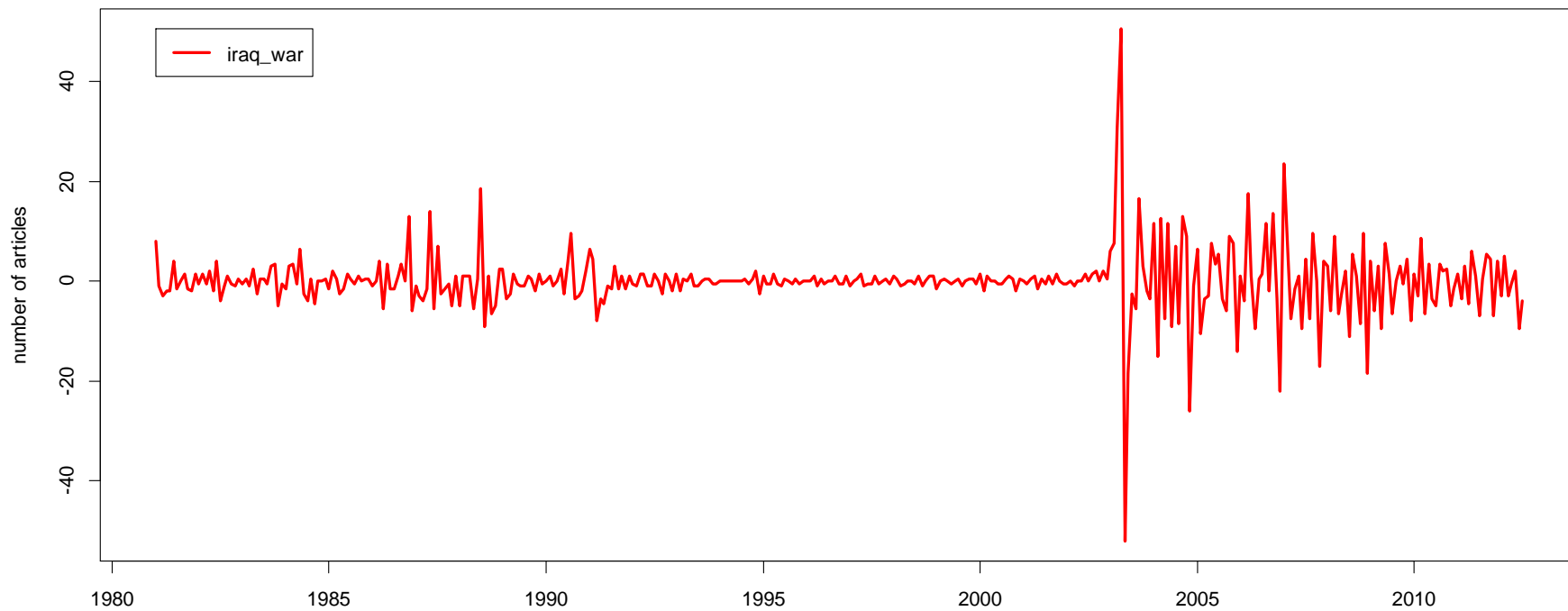


day-scale aggregation
Time Distribution of Articles from the New York Times



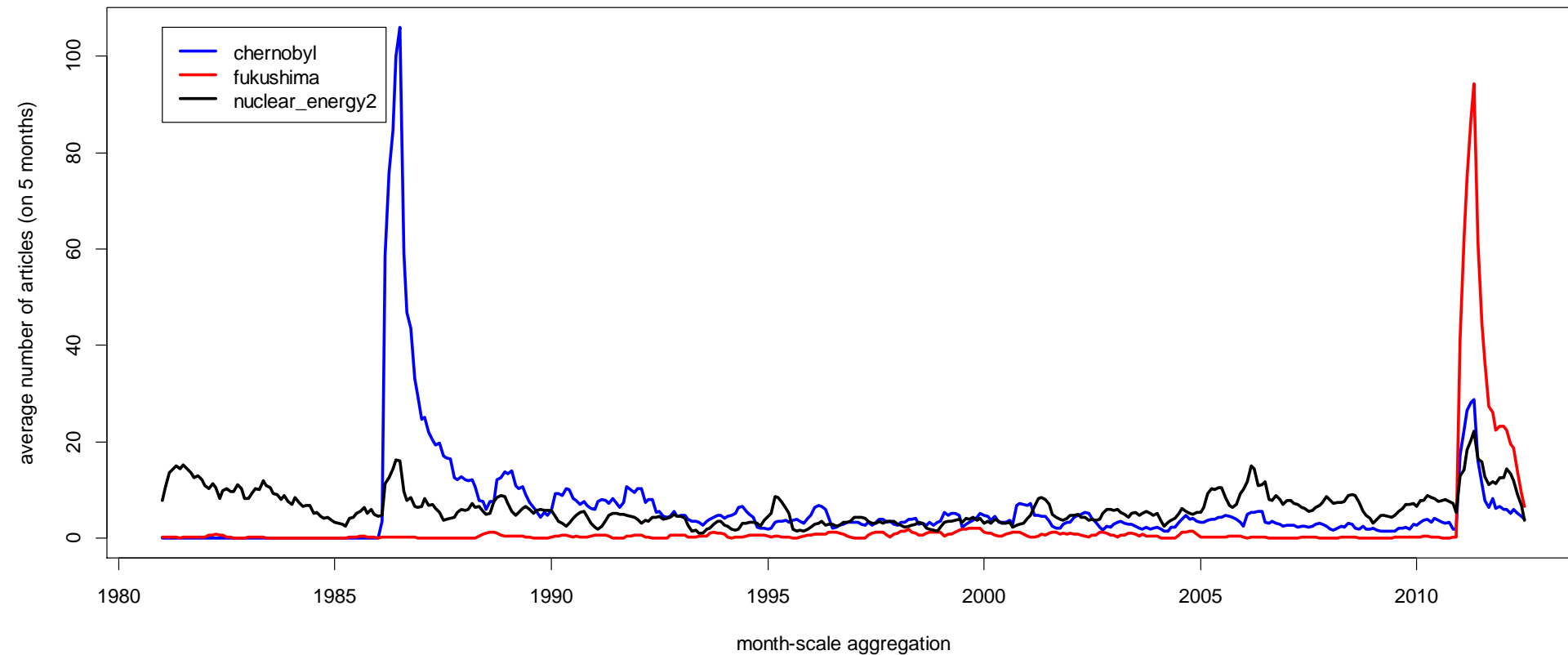


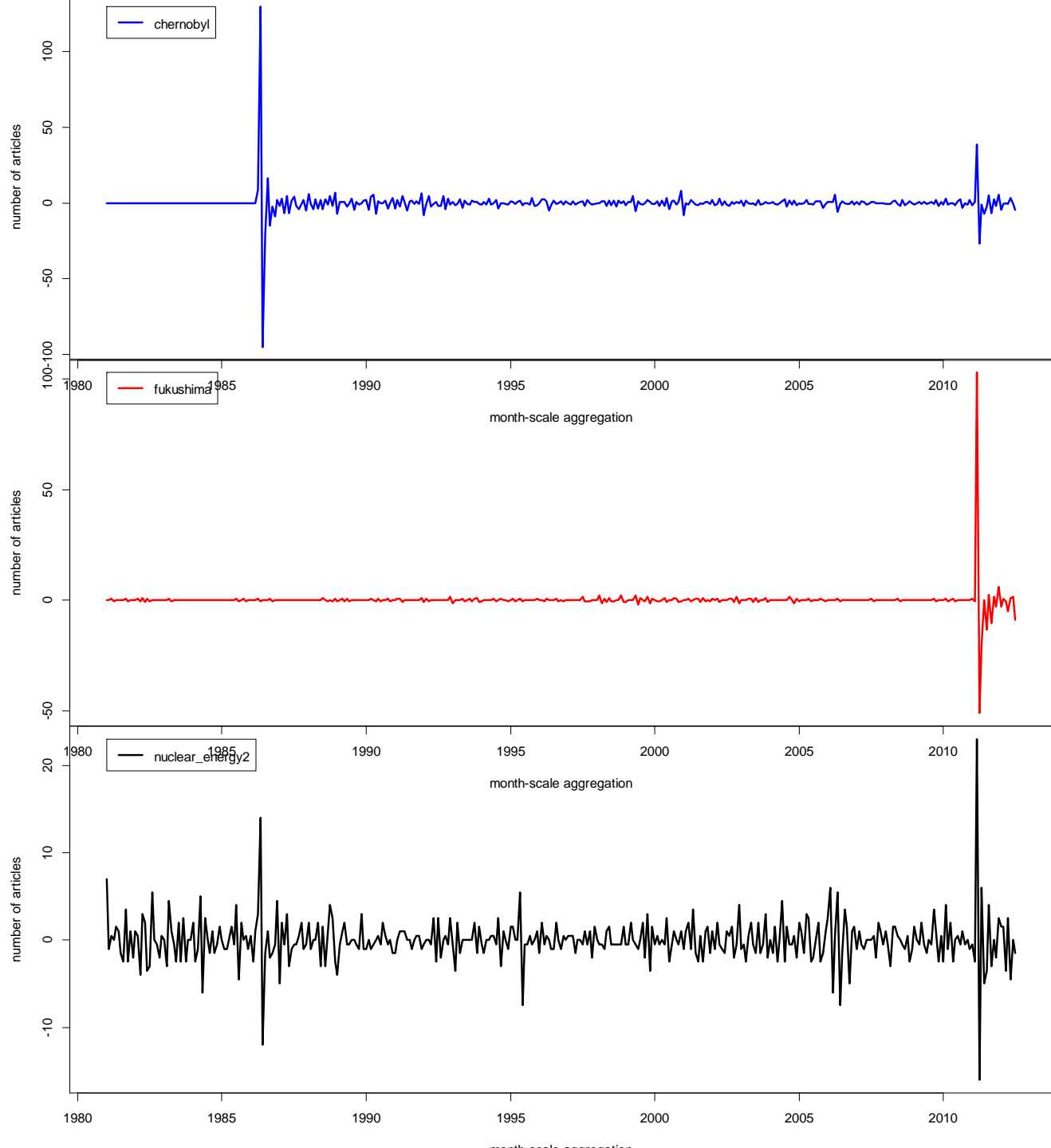
Time Distribution of Articles from the New York Times



Time analysis

Time Distribution of Articles from the New York Times





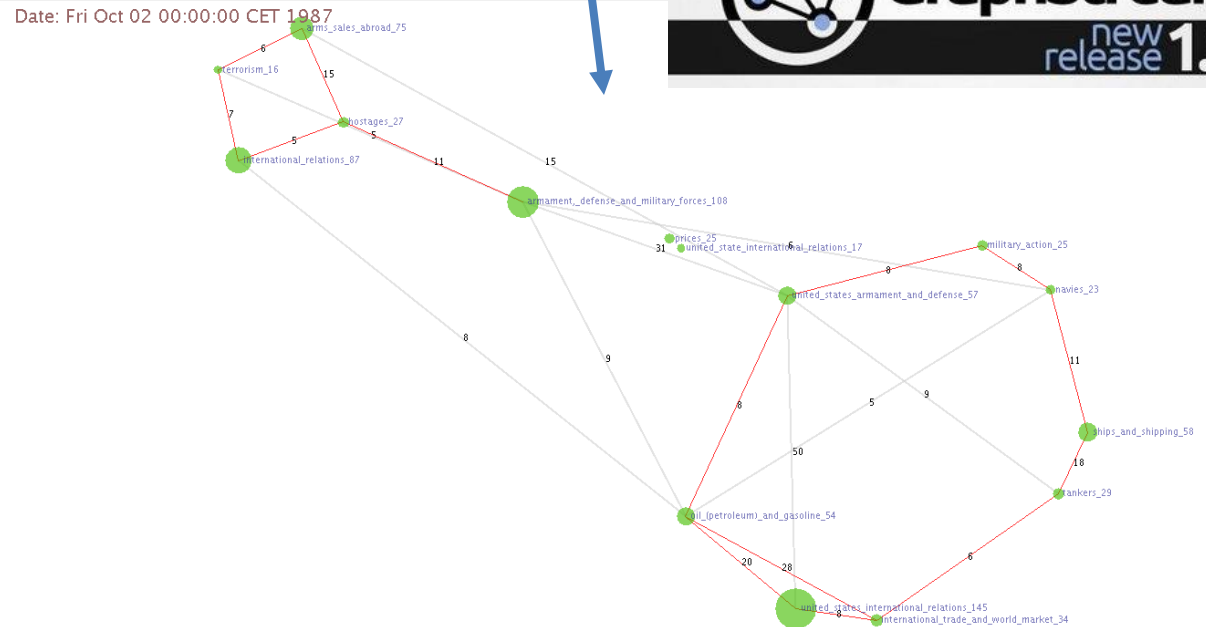
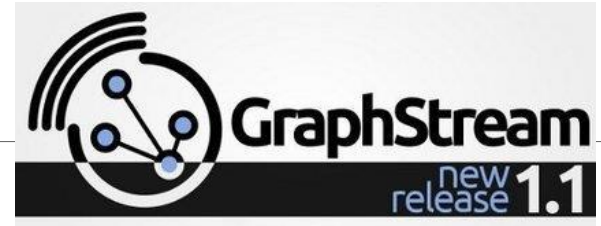
Data Retrieval

Fukushima

Chernobyl

Nuclear
energy

Topics from
NY Times



DEMONSTRATION

Thank you for your attention

Davide Beraldo

Juan Galán Páez

Robin Lamarche-Perrin