



DEFENCE AND SPACE

Application of AI Techniques to Deep Web Social Network Analysis

An Airbus & Litis work

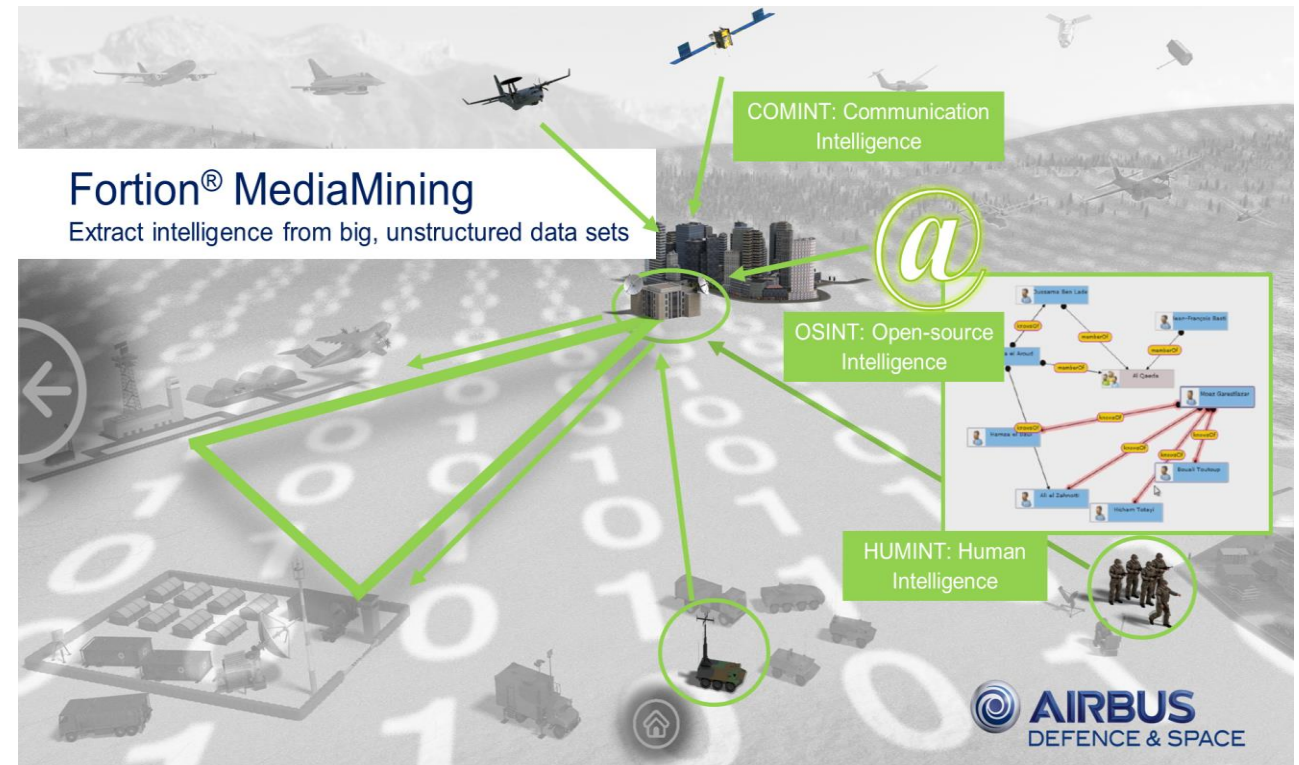
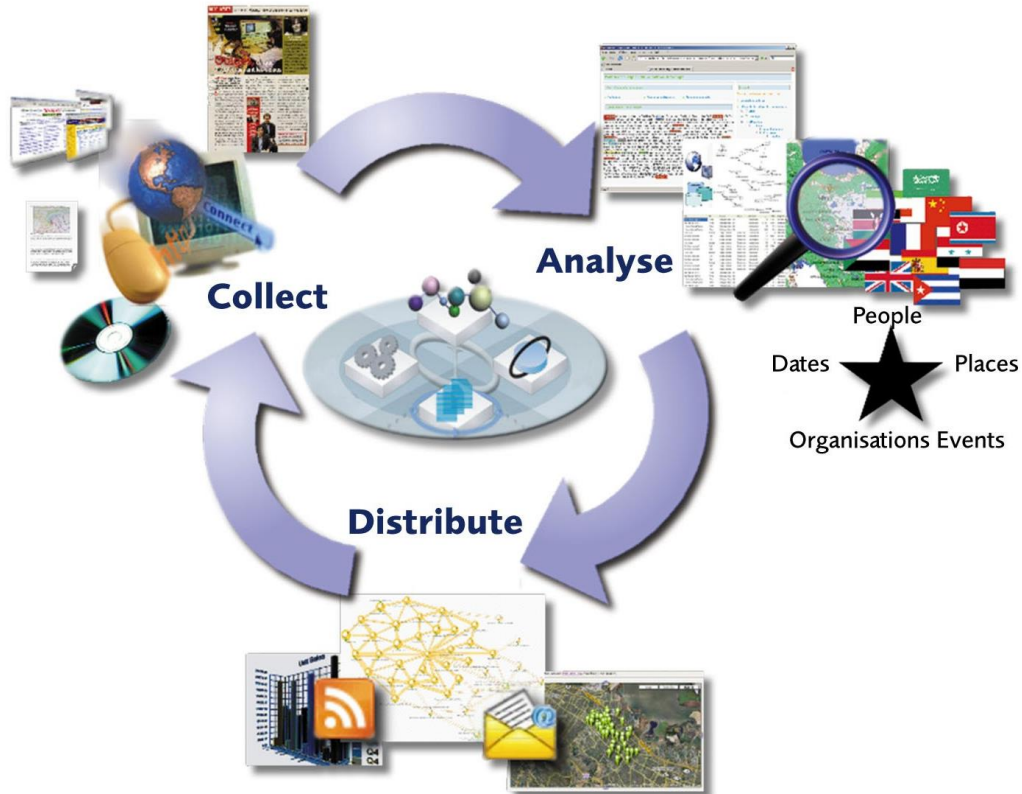


Guillaume GADEK, Stéphan BRUNESSAUX and Alexandre PAUCHET
31st MAY 2018

AIRBUS

Artificial Intelligence at AIRBUS Defence & Space: Open Source Intelligence

Fortion® Media Mining



Three steps of analysis for Social Media Intelligence

Text analysis

Sentiment, topic

Behaviour description

Hashtags, mentions, temporality

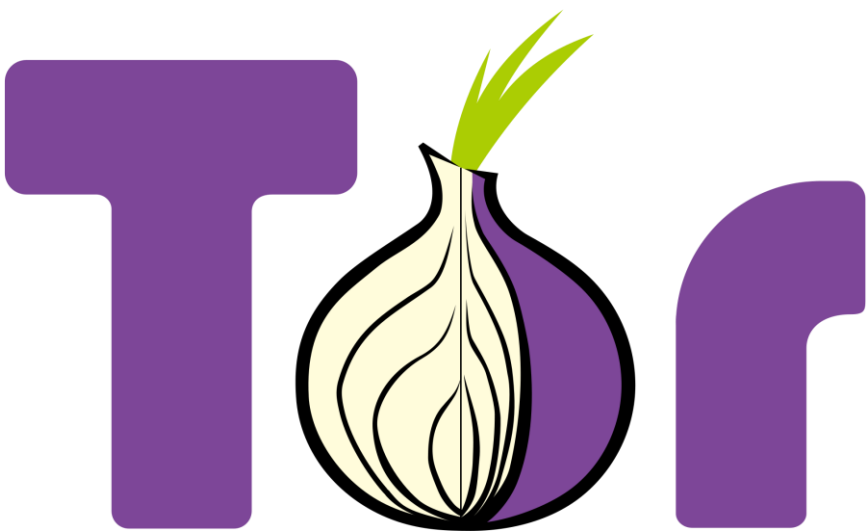
Group detection

Interaction-based communities



Today's target : Galaxy2, on TOR

TOR : The Onion Network



The anonymous Internet

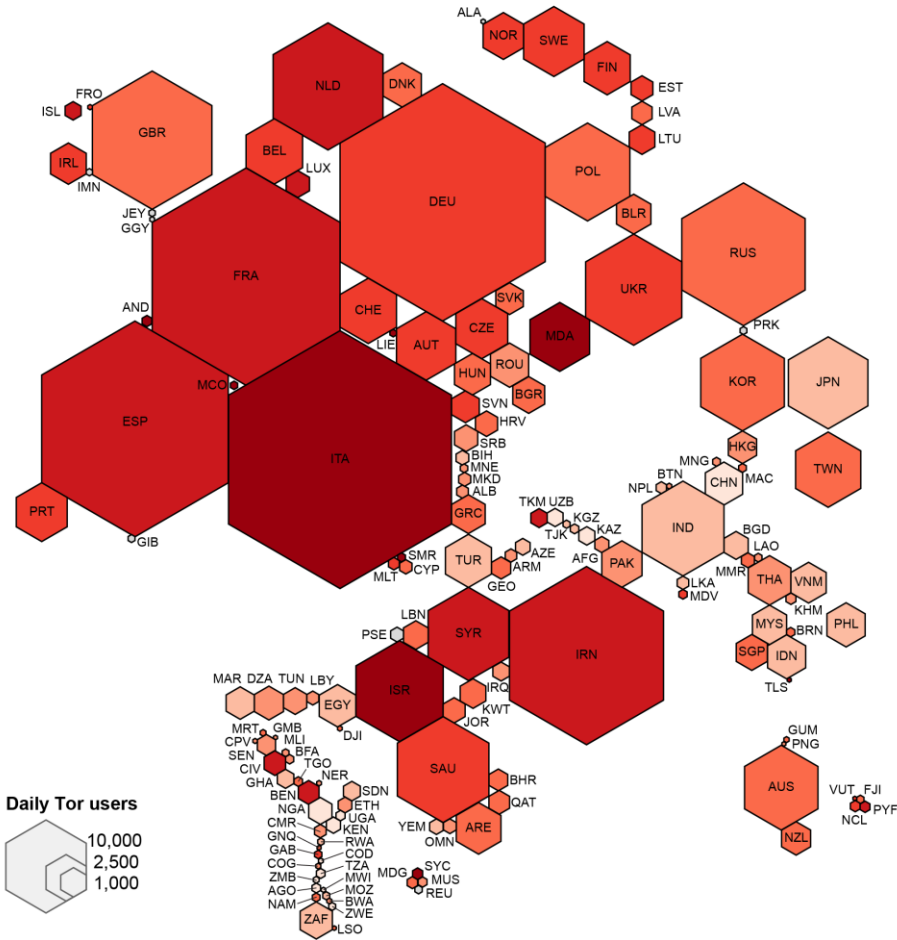
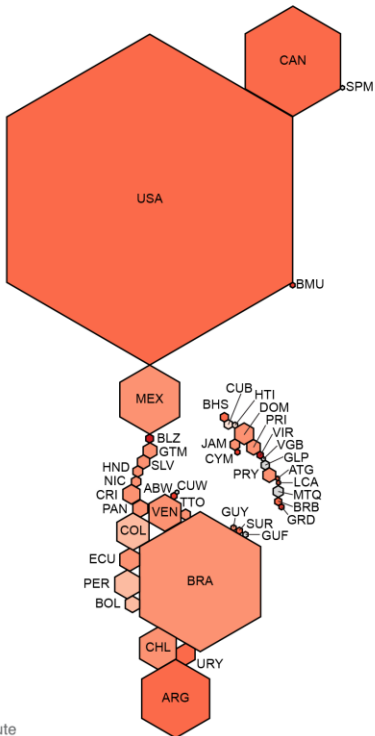
Daily Tor users
per 100,000
Internet users

- > 200
- 100 - 200
- 50 - 100
- 25 - 50
- 10 - 25
- 5 - 10
- < 5
- no information

Average number of
Tor users per day
calculated between
August 2012 and
July 2013

data sources:
Tor Metrics Portal
metrics.torproject.org
World Bank
data.worldbank.org

by Mark Graham
(@geoplace) and
Stefano De Sabbata
(@maps4thought)
Internet Geographies at
the Oxford Internet Institute
2014 • geography.oii.ox.ac.uk



Today's target : Galaxy2, on TOR

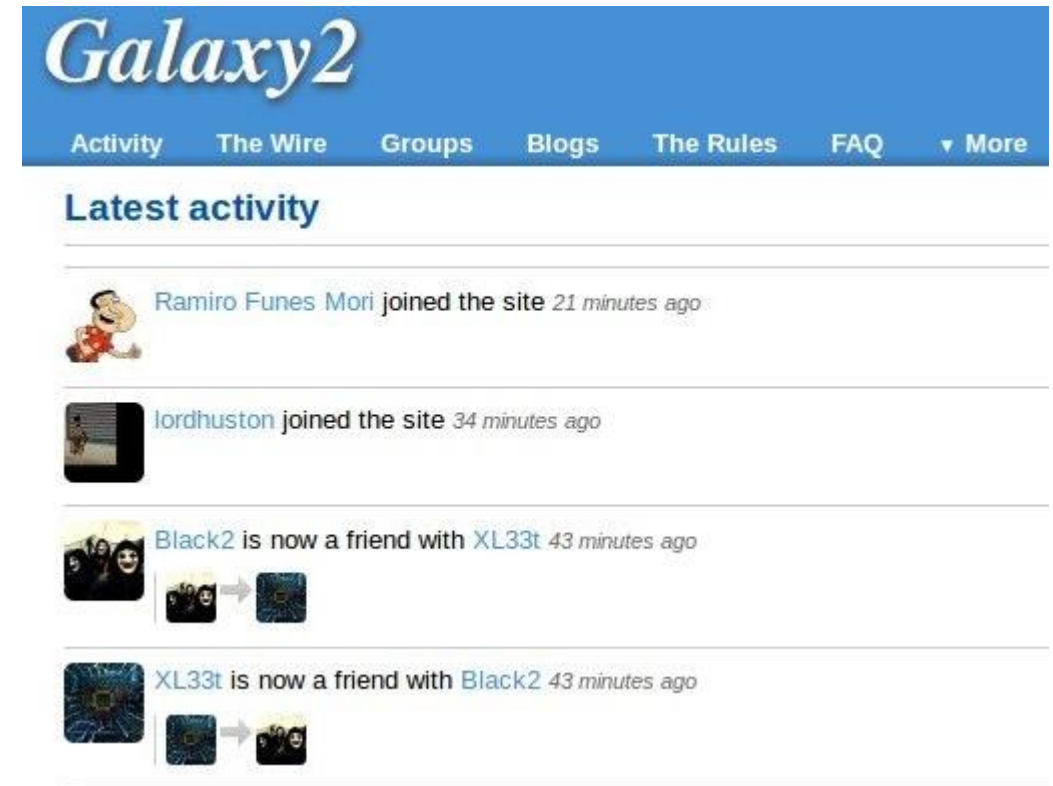
Galaxy2: "the most popular social network on TOR"

Active during 2015-2016-2017, disrupted since.

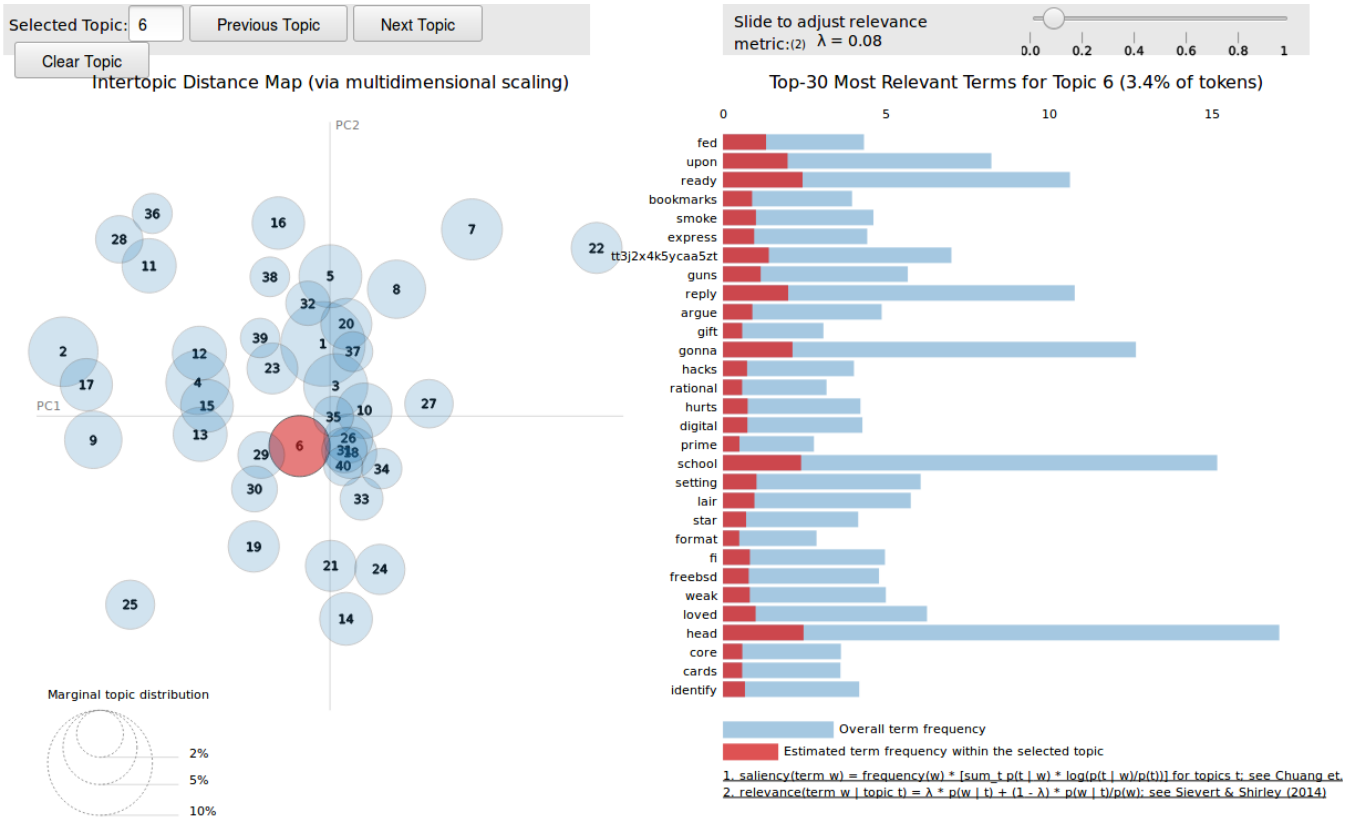
Based on the **elgg** open source framework.

Microblogging and friendship features.

About **20,000 users** in total.

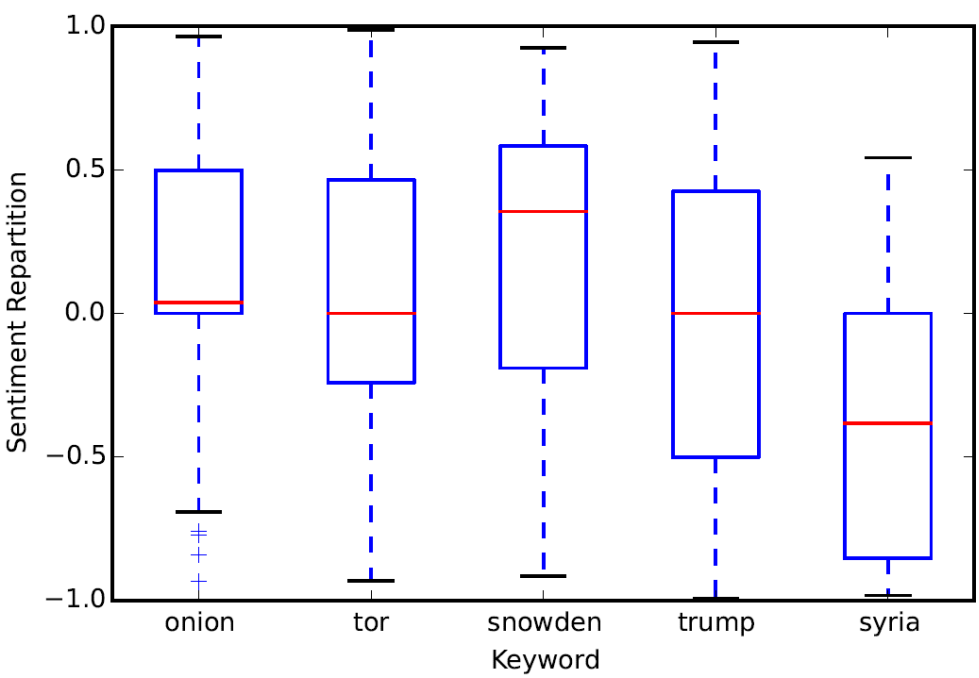


Textual Analysis: Topic and Sentiment



Emergent topic detection and description [C. Sievert 2014]

Coupling Sentiment [C.J. Hutto 2014] with keywords



User profiling

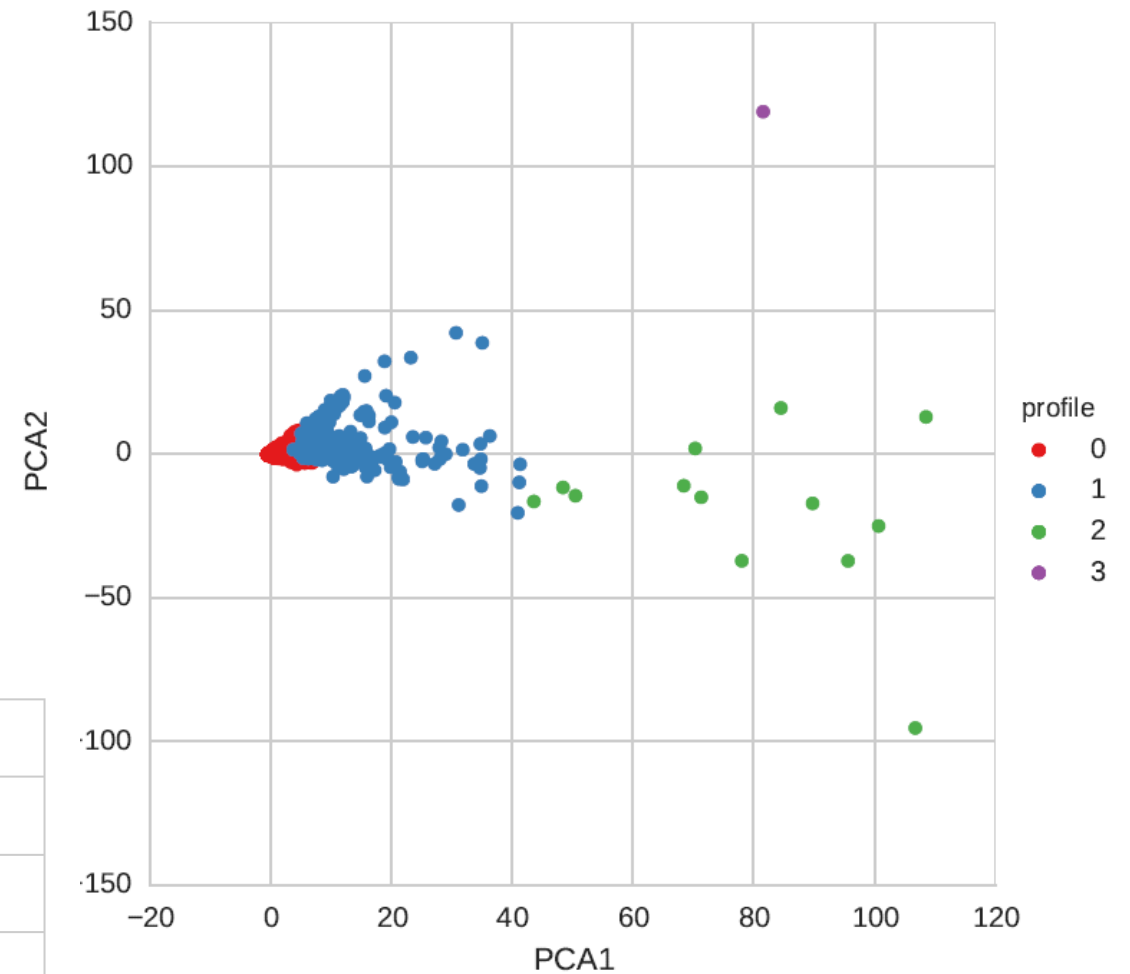
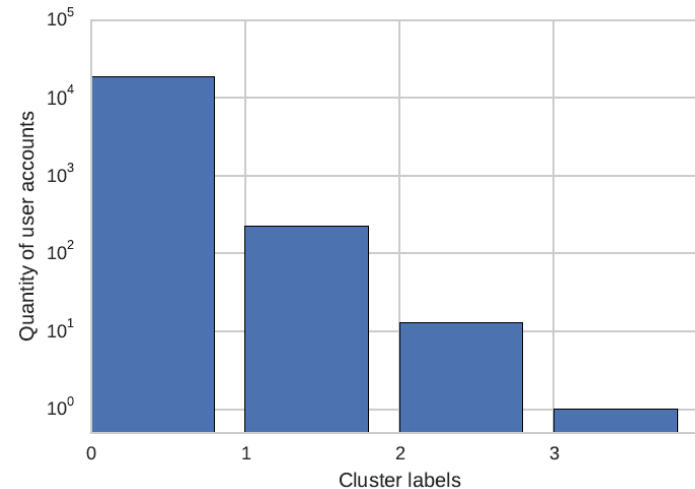
Machine-learning based profiling

Unsupervised types of behaviours

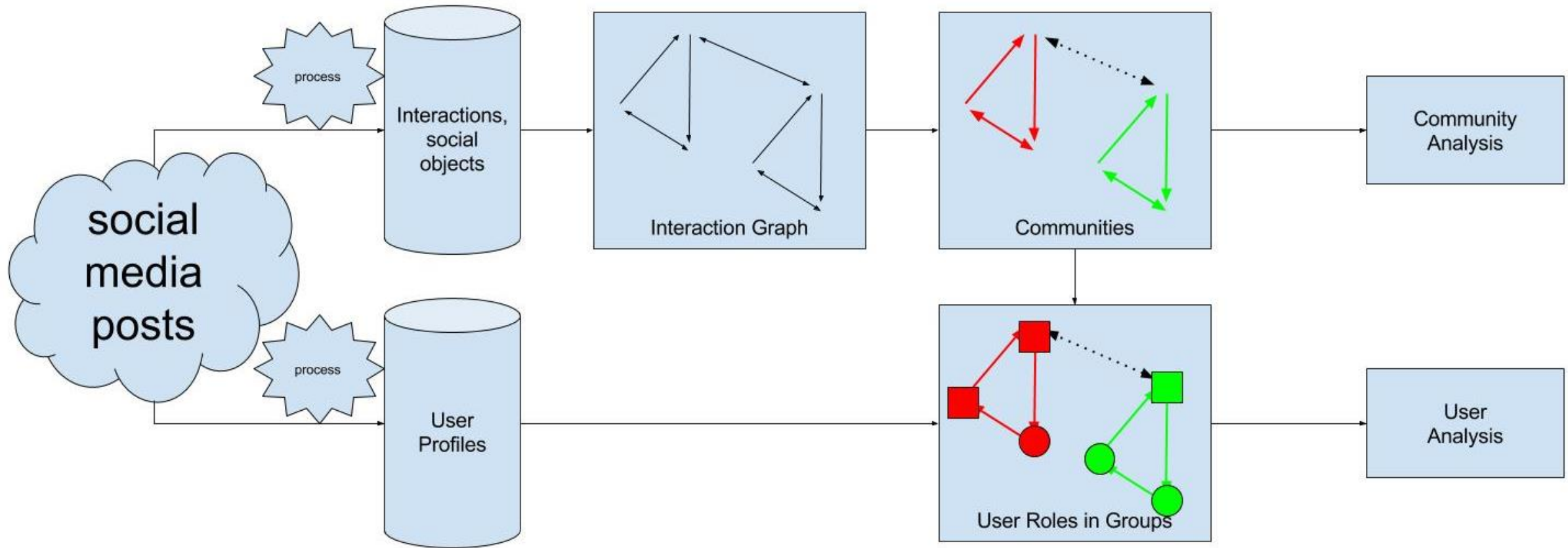
Statistical study of specific user, aspect by aspect

Steps:

1. Data selection
2. Data cleaning
3. Dimension Reduction
4. Unsupervised Clustering
5. Explanation and Exploitation

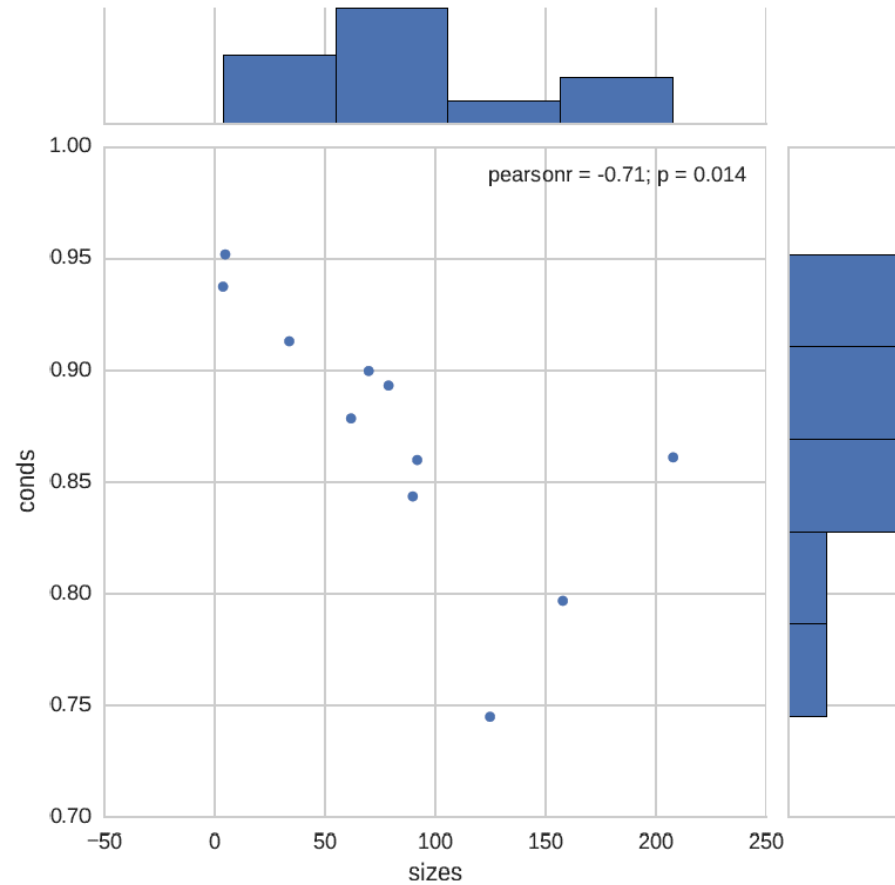


Groups: the social dimension, represented by graphs



Groups: the social dimension

Detection and Characterisation

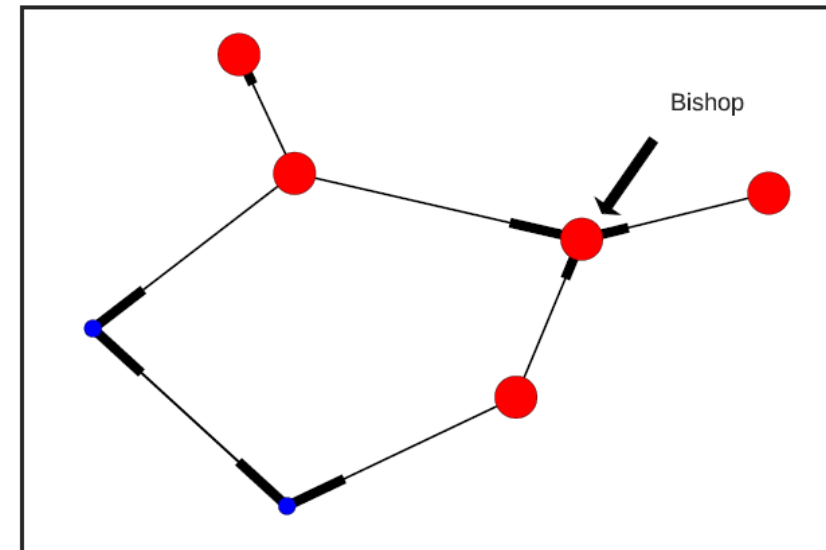


Looking for groups with:

- Strong **interaction**
- Common **topics of interest**

Exploring a graph of interaction between user accounts

Zooming in a specific group



Conclusion: AI needs in SocMInt

- **Three levels of analysis:**
 - Textual: sentiment, emotion, topic
 - User: profiling, behaviours, influence
 - Groups: detection, impact, link strength, exploration
- New challenges, and new requests
- Task-specific modules to benchmark and integrate in a larger solution
- General difficulties to obtain the data
 - GPDR & privacy
 - Proprietary data, access limitations
- Some NATO work on the domain:
 - NIAG study on Big Data, usecase on social media analysis
 - StratCom CoE: excellent publications such as *“Robotrolling”*

Thank you