**Project Ideas**

Panel data on networks: Data Source GDELT project

*What does the GDELT project offer? – GDELT Global Knowledge Graph (GKG)*

“…connects every person, organization, location, count, theme, news source, and event across the planet into a single massive network that captures what’s happening around the world, what its context is and who’s involved, and how the world is feeling about it, every single day.”

* Links in a potential network: connection in the press -> co-occurrence of the people/topics/locations… (whatever we want as our nodes) in same articles during a specific period of time.
* It is possible to make a query that records the articles each pair of names appeared together.
* Information on links: all information about the article (date, location, tone, content…)
* Information on nodes???
* List of topics-classification in GKG: 

Easy to download data:

<http://analysis.gdeltproject.org/module-event-timeline.html> (event times)

<http://analysis.gdeltproject.org/module-gkg-wordcloud.html> (word clouds)

<http://analysis.gdeltproject.org/module-gkg-network.html> (networks)

<http://analysis.gdeltproject.org/module-event-timemapper.html> (maps)

<http://analysis.gdeltproject.org/module-gkg-heatmapper.html> (heatmap)

Possible applications (very much idea-stage):

1. Find sth that is talked about first separately at different places in the world, capture their local networks and then maybe see how it becomes a global network.
2. Find an interesting topic/person and download its network at different points in time to see how it evolved. We could use this for contagion analysis since we have information behind the content of connection between our nodes.

**Kaggle DATASET – project ideas:**

1. <https://www.kaggle.com/ssalgadom/k12-ecuador-tuition-and-enrollment-fees-20182019>: Ecuador tuition fees in primary and secondary schools
2. <https://www.kaggle.com/mzwelch/national-education-longitudinal-survery>: Longitudional Survey for Education
3. <https://www.kaggle.com/jboysen/us-perm-visas>: US VISA for permanent stay - network analysis, we have employer and possible employee info - possible network analysis
4. <https://www.kaggle.com/fifthtribe/how-isis-uses-twitter>:

Social Network Cluster Analysis, Keyword analysis, Data categorization of links, sentiment analysis, timeline view

Use of network and text data

1. <https://www.kaggle.com/geocomuna/rural-schools-mexico-2016-sep-snie>: School data in developing countries - spatial analysis and network analysis, for example here Mexico
2. <https://www.kaggle.com/crowdflower/political-social-media-posts>: Political social media posts in the US: text analysis and network analysis, facebook vs. twitter
3. <https://www.kaggle.com/reubencpereira/spatial-data-repo>: Regional level data in developing countries (environmental, climate, socioeconomic/demographic data, conflict data, natural disaster data) - machine learning, spatial data, network analysis - <https://www.kaggle.com/brookewatson/worldbank-subnational-poverty> (further regional poverty data)

<https://www.senate.gov/legislative/votes.htm>

US legistlative votes. Using machine learning to figure out on what topics senators divert from their party/the president.