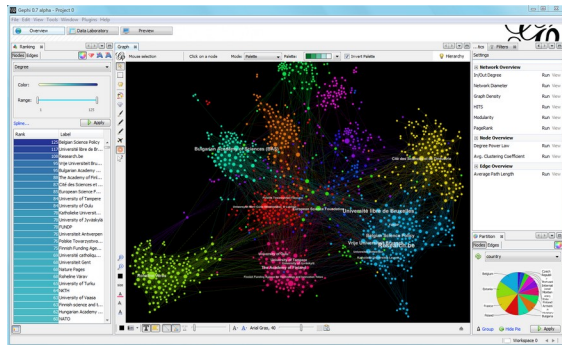


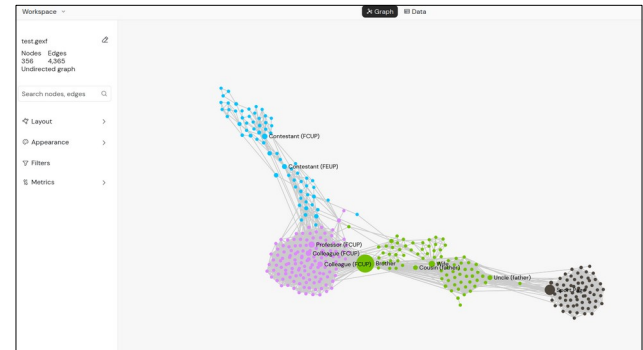
# Network Analysis and Visualization with Gephi



**Pedro Ribeiro**  
(DCC/FCUP & CRACS/INESC-TEC)



*Gephi* Gephi



## Why Visualization?

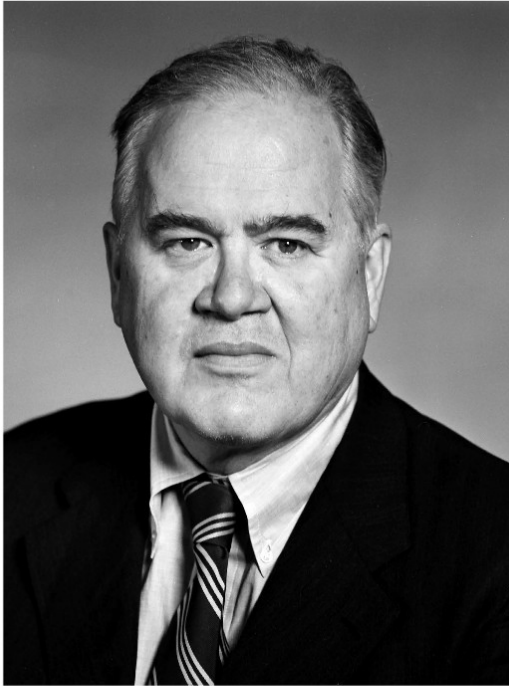


Photo Credit: Princeton University, Robert Matthews

“The greatest value of a picture is when it forces to notice what we never expected to see”

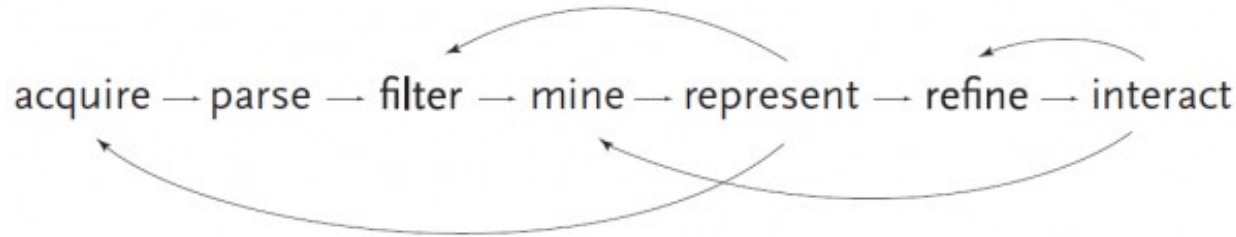
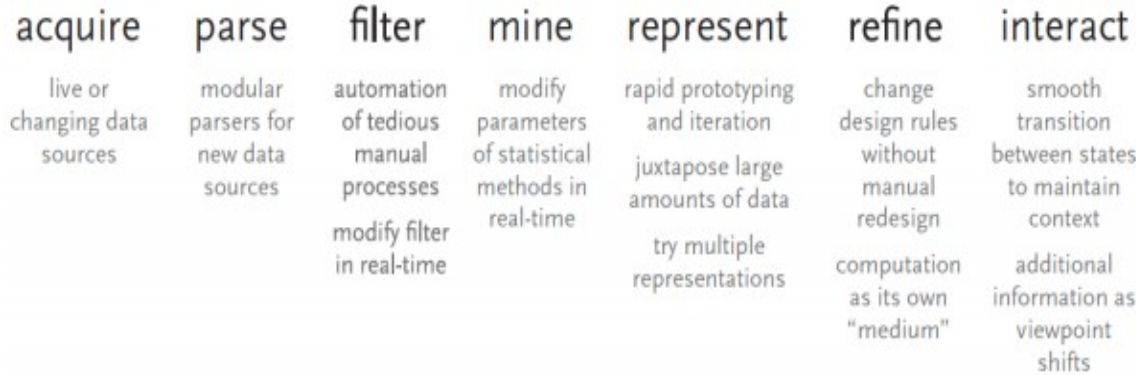
*John W. Tukey*

# Exploratory Data Analysis

## Visualization alone is not enough:

- Part of a larger process to extract insight

## Data Process Chain:

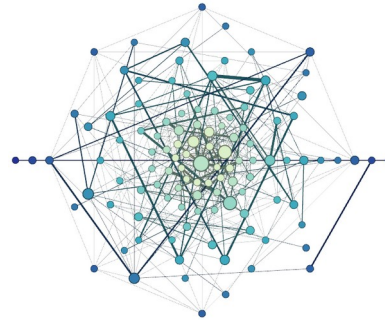
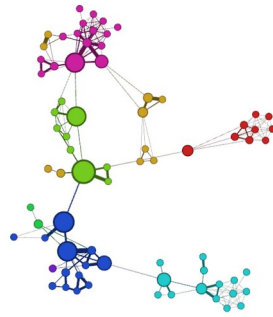


**Non-linear  
Trial and Error!**

# Exploring a Network

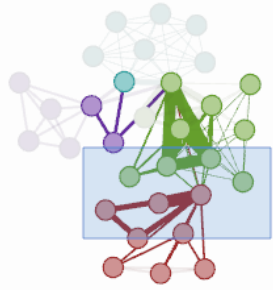
## 1) See the network

- Draw using a certain layout, ...



## 2) Data Process Chain

- Group, filter, compute metrics, ...



## 3) Build a visual language

- Size of nodes, thickness of edges, colors, ...

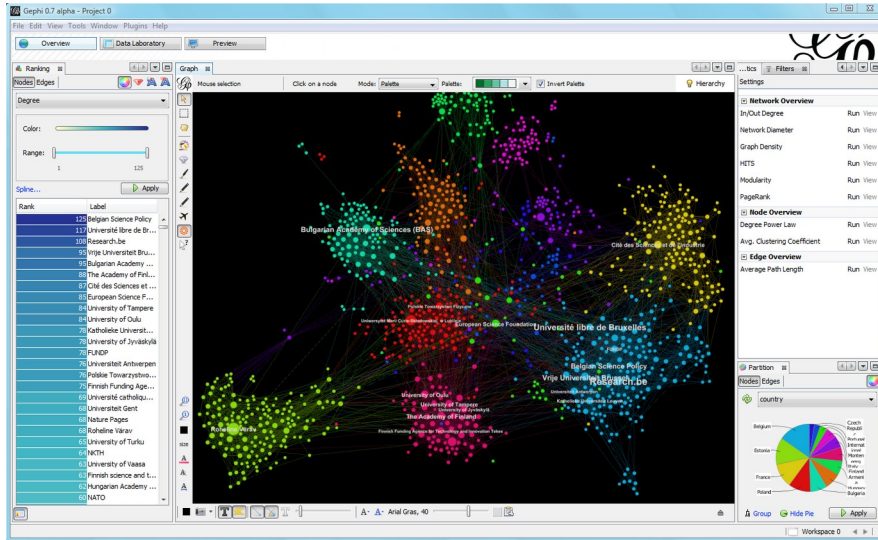


# Exploring a Network

Today we are going to use



- Open-Source Network Analysis and Visualization Platform (written in Java)



Like Photoshop™ / GIMP / Inkscape  
for graphs

The goal is to help to make hypothesis and intuitively discover patterns.

It is a complementary tool, as visual thinking with interactive interfaces can facilitate reasoning.

# Why Gephi?



## Because it is available for different platforms

- Windows / MacOS / Linux / Browser

## Because it has a large community

## Because it has history (and will continue to have)

- Started at 1998 / Maintained by a consortium (long-term vision)

## Because it is extensible with plugins

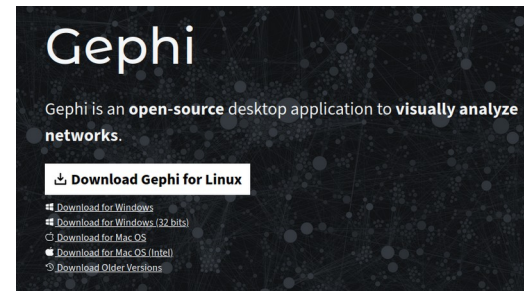
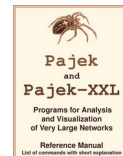
- Gephi marketplace

### Limitations:

- Still in beta version so there are a few rough edges
- Not prepared to handle very large networks (depending on the infrastructure – RAM mostly - can manage networks up to 100.000+ nodes)

## There are other options

- The main concepts and ideas we will show can be used on any other visualization tool



# Network Analysis with Gephi

## Goal of today's lecture:

Consolidate the main concepts and techniques learned in the previous lectures by performing a (*hands-on*) network analysis of several real world networks

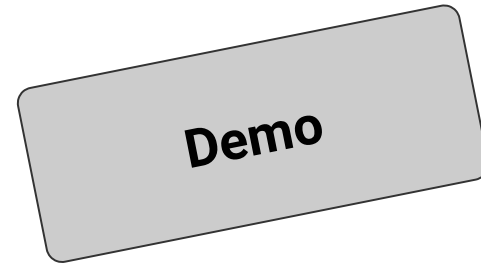
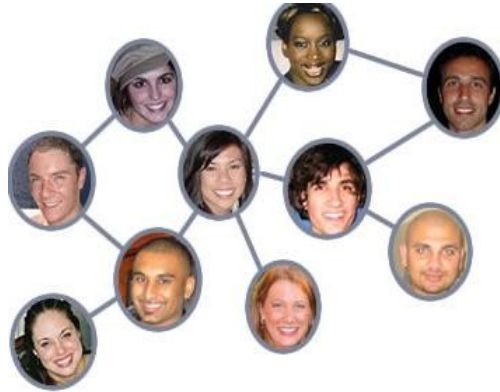
### Specific Goals:

- Perform an empirical analysis of the network
- Loading Networks (*opening, importing raw data, ..*)
- Computing Metrics (*centralities, degrees, distances, communities, ...*)
- Filtering (*main operators, selecting, ranges, combining, ..*)
- Create a clear and simple to understand visualisation of the network (*color or size of the nodes and edges according to a metric or partition, ...*)

# Network Analysis with Gephi: Facebook Network

## My own facebook (ego) network - `facebook.gphi`

- Nodes are users and links represent *friendships* (*undirected graph*)
- Ego Network: all nodes connected to me and their connections (*without myself*)
- Collected automatically (*there used to be plugins for that*)
- 356 persons, 4,365 connections



# Network Analysis with Gephi: Flights Network

Flights Data (OpenFlights) – `airports.csv` and `routes.csv`

- <https://openflights.org/data.php>
- Compiled (also) by Open Flights website users
- 3,154 airports, 66,500 routes from 538 airlines
- Made for showing GeoLayout

OpenFlights.org



Demo

# Network Analysis with Gephi: Languages Network

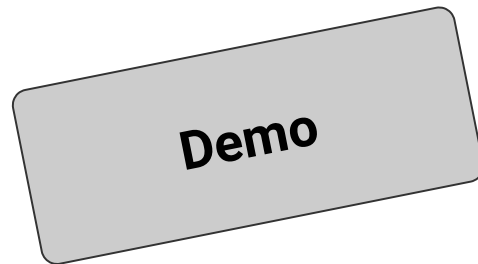
Countries and Languages – `unicode_nodes.csv` and `unicode_edges.csv`

- Unicode CLDR Project

[https://www.unicode.org/cldr/charts/47/supplemental/territory\\_language\\_information.html](https://www.unicode.org/cldr/charts/47/supplemental/territory_language_information.html)

- Made for showing projections of a bipartite network

## Unicode CLDR Project



## Getting started with Gephi: Some Resources



- [Gephi.org](http://gephi.org) (*homepage*)
  - [Learn how to use Gephi](#) (*several tutorials available*)
  - [Gephi Tutorial](#) by Martin Grandjean
  - [Gephi Network Datasets](#)
  - [Gephi Plugins](#)
- **Gephi Video Tutorial:** <https://www.youtube.com/watch?v=rnCTpzY2xUM>  
(*from myself from a previous Network Science course*)



**GEPHI**

**NETWORKS  
SCIENCE  
STUDENTS**

**IMPLEMENTING  
CODE**