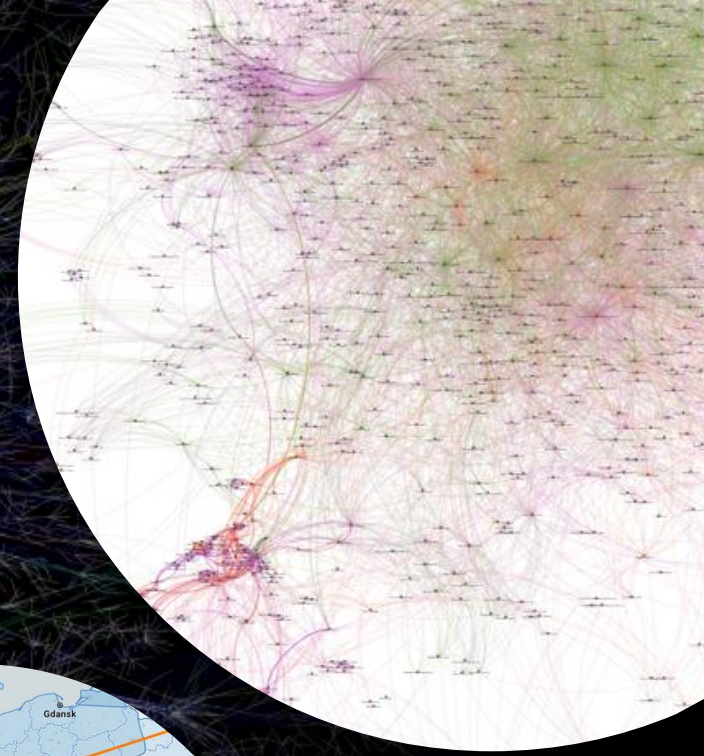
EnExDi 2025

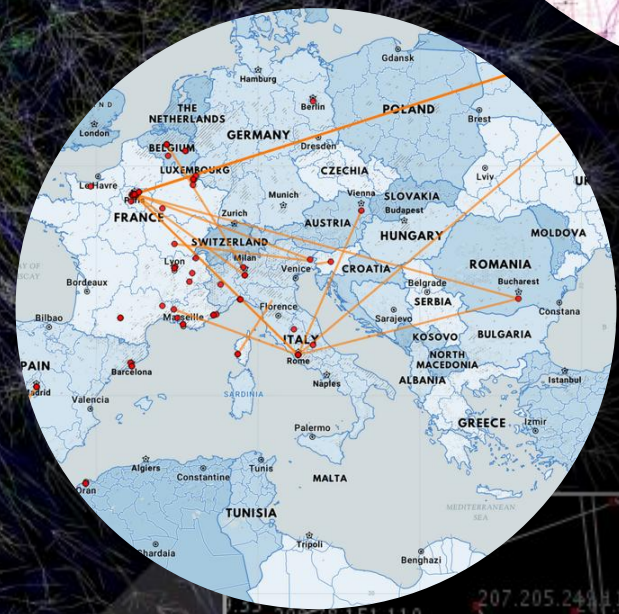
Network Analysis

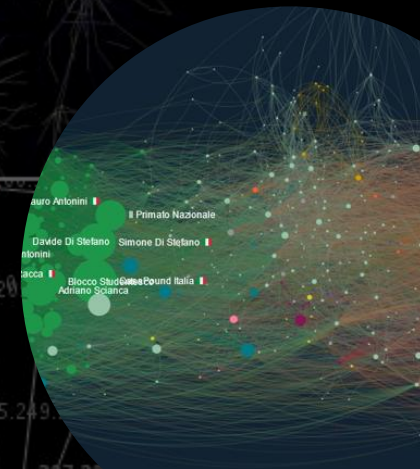
Martin Nicastro – nicastromartin@gmail.com

Musicology PhD (University of Pavia)

Researcher and R coder (Live Music Mapping Project)



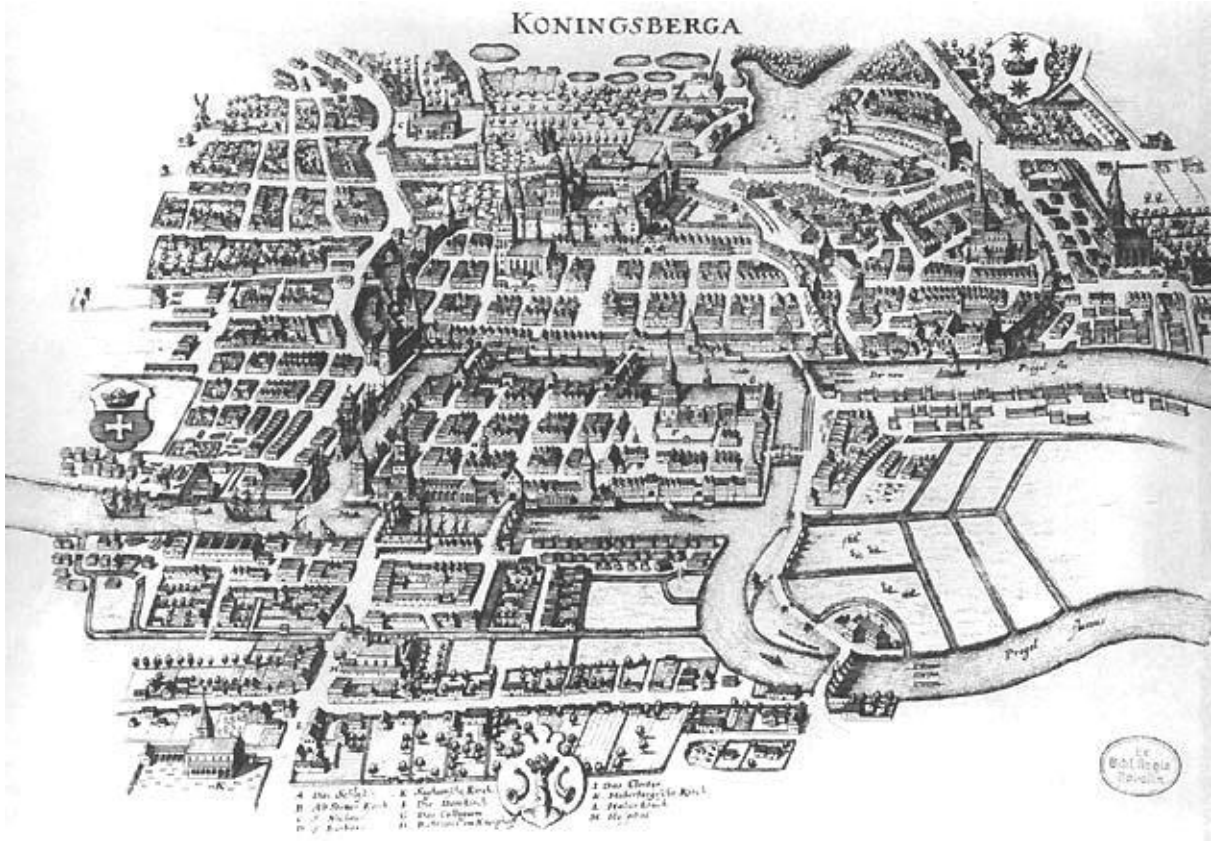




Graph theory

- The origins of graph theory date to the resolution of the **mathematical problem of the "Seven Bridges of Königsberg"** by Leonhard Euler in 1736.

- Is it possible to cross all seven Königsberg bridges in one route, i.e. without ever using the same bridge twice?



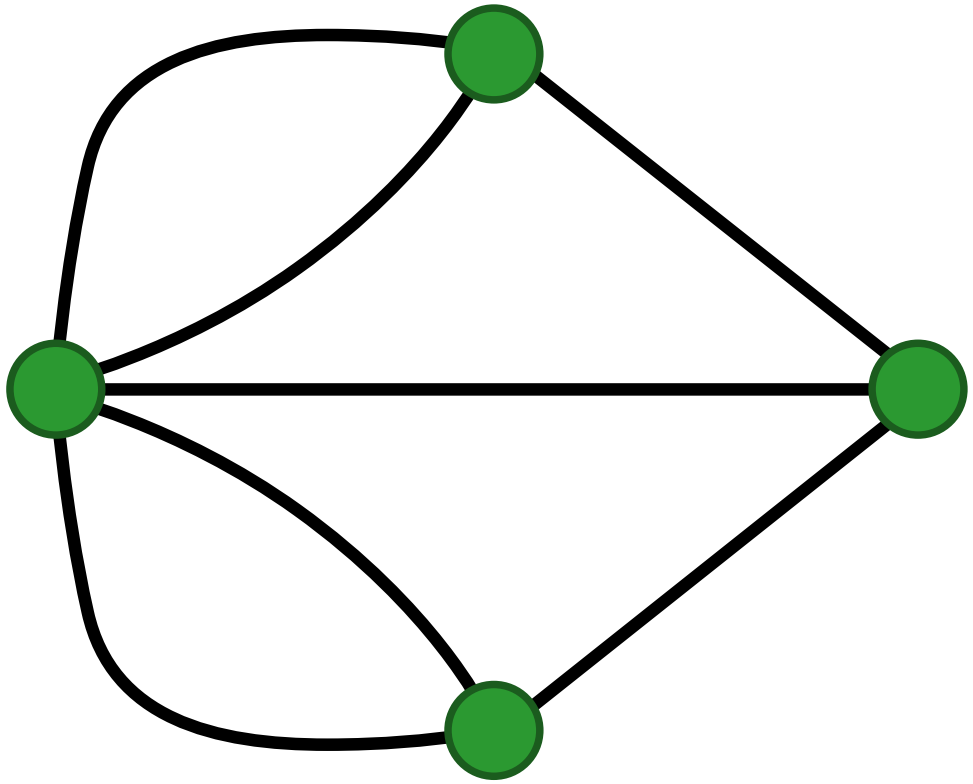
Merian-Erben, Public domain, via Wikimedia Commons

Graph theory

- Euler's proof is based on a **process of graphical abstraction**: each land mass becomes a point, each bridge a line.

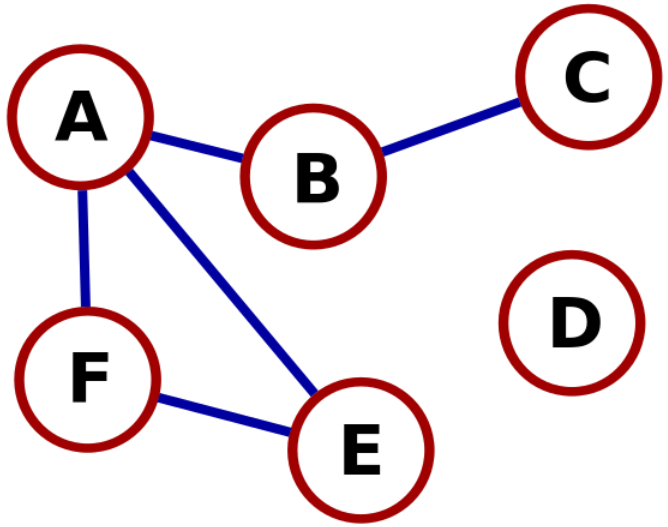
- Points with an odd number of connections must necessarily be start or the end of the route. A continuous path can only have one starting and ending point: **the seven bridges cannot be crossed without retracing one's steps.**

- **The solution to the problem is inherent in the structure of the network**, whose properties define its limits and possibilities



Graph theory

- The definition of **"graph"** means a relational structure made up of a finite number of vertices (or **nodes**) and a finite number of segments (called **edges**).

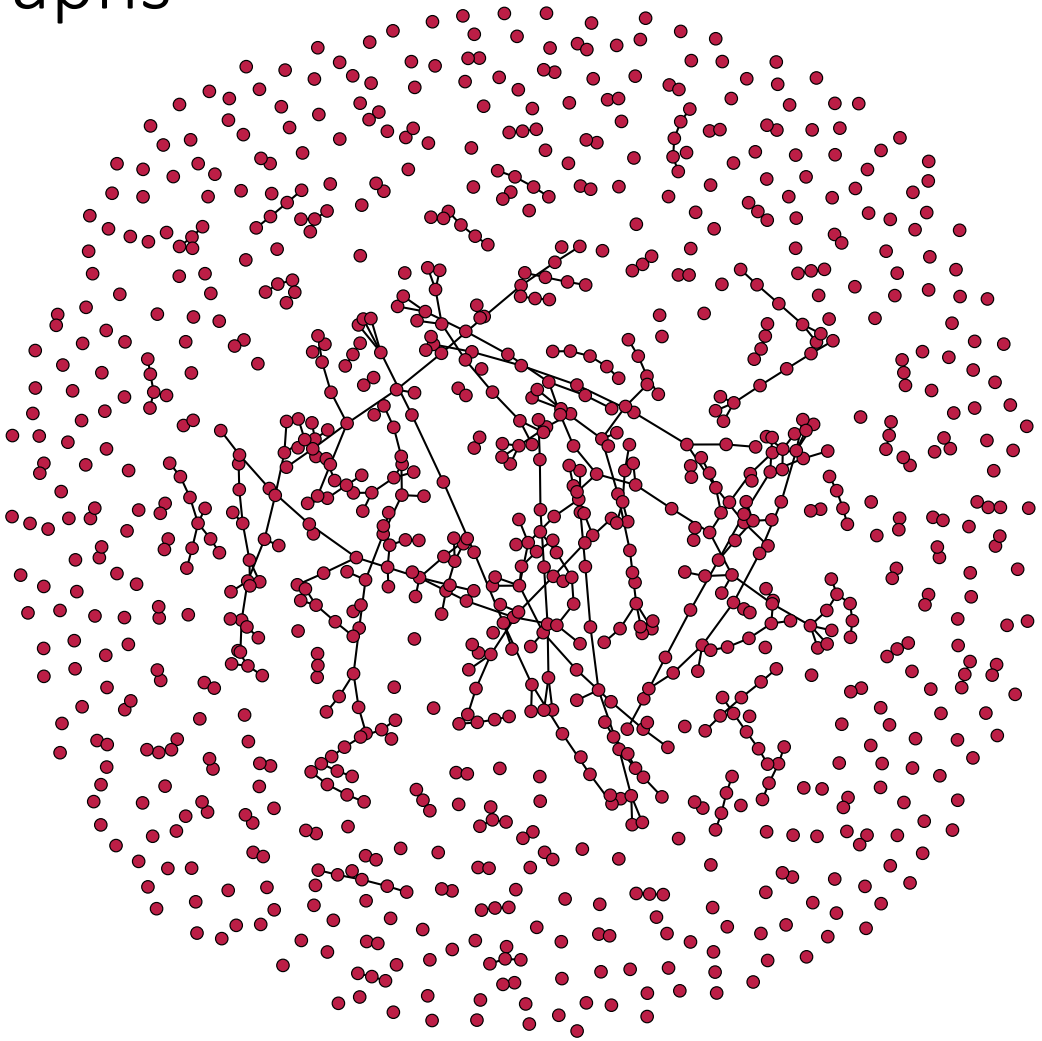


The Erdős–Rényi model: random graphs

- Hungarian mathematicians Alfréd Rényi and Paul Erdős were among the first to adventure beyond the description of the properties of graphs. Their model, introduced in 1959, investigates the **modalities of their formation**.

- Starting from the assumption that each node has equal possibilities of developing a relationship, Rényi and Erdős demonstrated how **for the formation of a single "component" it is sufficient for each node to have a single bond.**

- By **"component"** we mean a network in which it is possible to move from one node to another through a more or less large number of connections

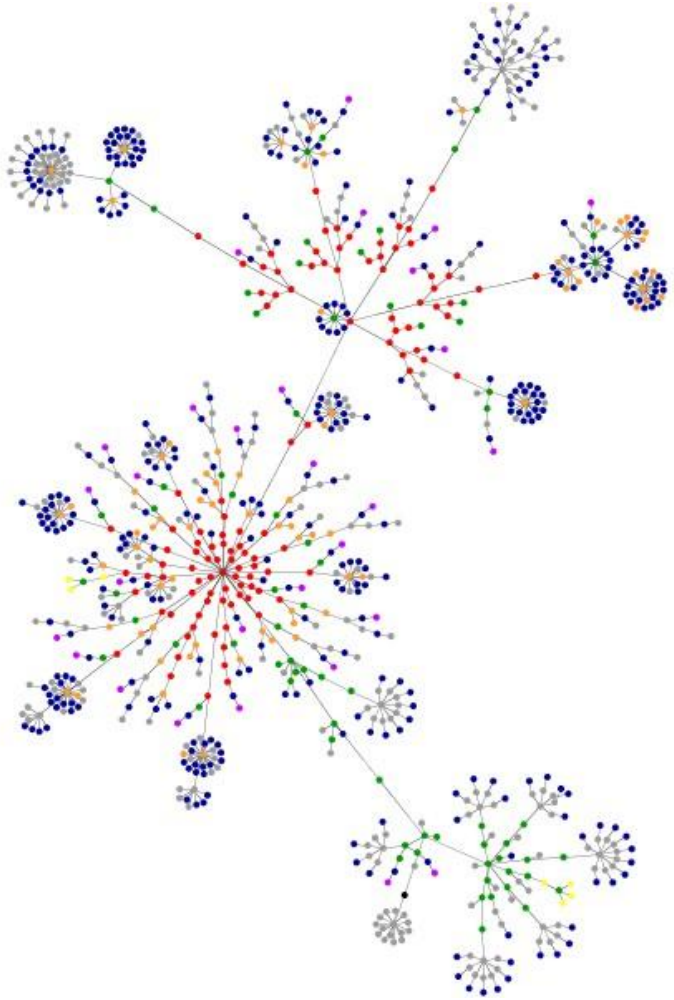


By David Eppstein - Own work, CC0, https://commons.wikimedia.org/w/index.php?curid=115108260

The Barabási–Albert model: complex networks

- The experimental results obtained by Réka Albert, Hawoong Jeong and Alber-László Barabási in their **study on the reticular structure of the World Wide Web (1999)** profoundly question the random graph model: **80% of the links present on the Web it is directed at 15% of pages.**

- The study of **complex networks**, based on empirical data, highlights the existence of asymmetries: **some nodes have a much higher force of attraction than others** and catalyze a greater number of ties.

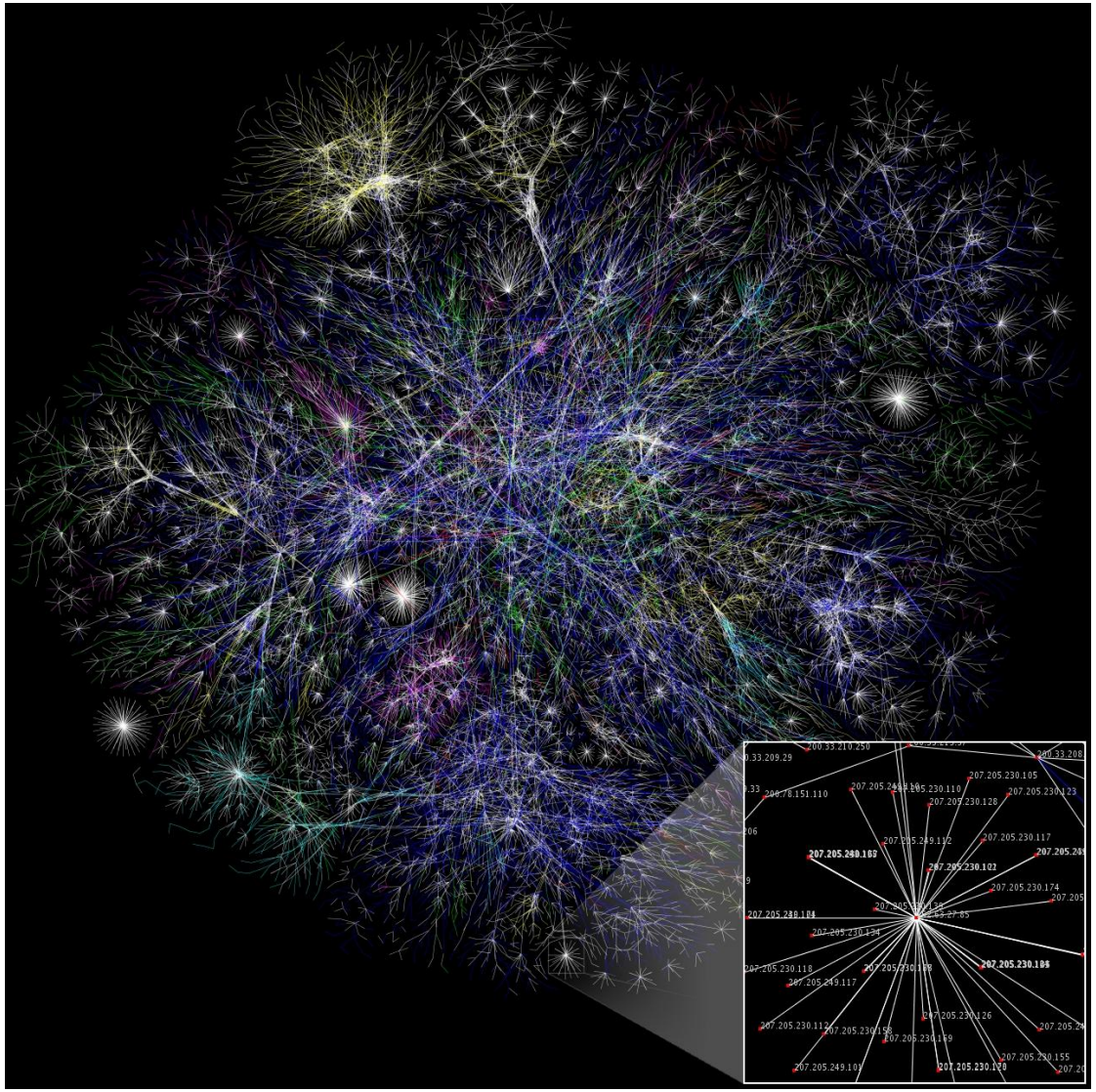


Sala (http://www.aharef.info), Public domain, da Wikimedia Commons

The Barabási–Albert model: complex networks

- Some nodes contribute more than others to the overall cohesion of the network, thus leading to the formation of “hubs". **The World Wide Web is structured around a few nodes with numerous links.**

- The Barabási–Albert (1999) model tries for the first time to account for this characteristic (absent in a random model for which each node has on average the same number of edges). **Its introduction broadened the range of applications of graph theory.**



Network Analysis

- "**Network theory**" is today applied in a vast variety of scientific fields, from particle physics to neuroscience, from the study of international trade to that of natural ecosystems. **Countless human activities and natural conformations can be described through the complex network model.**

- The study of complex networks allows us to investigate the **role of each node within the network** and **the overall structure of the network.**

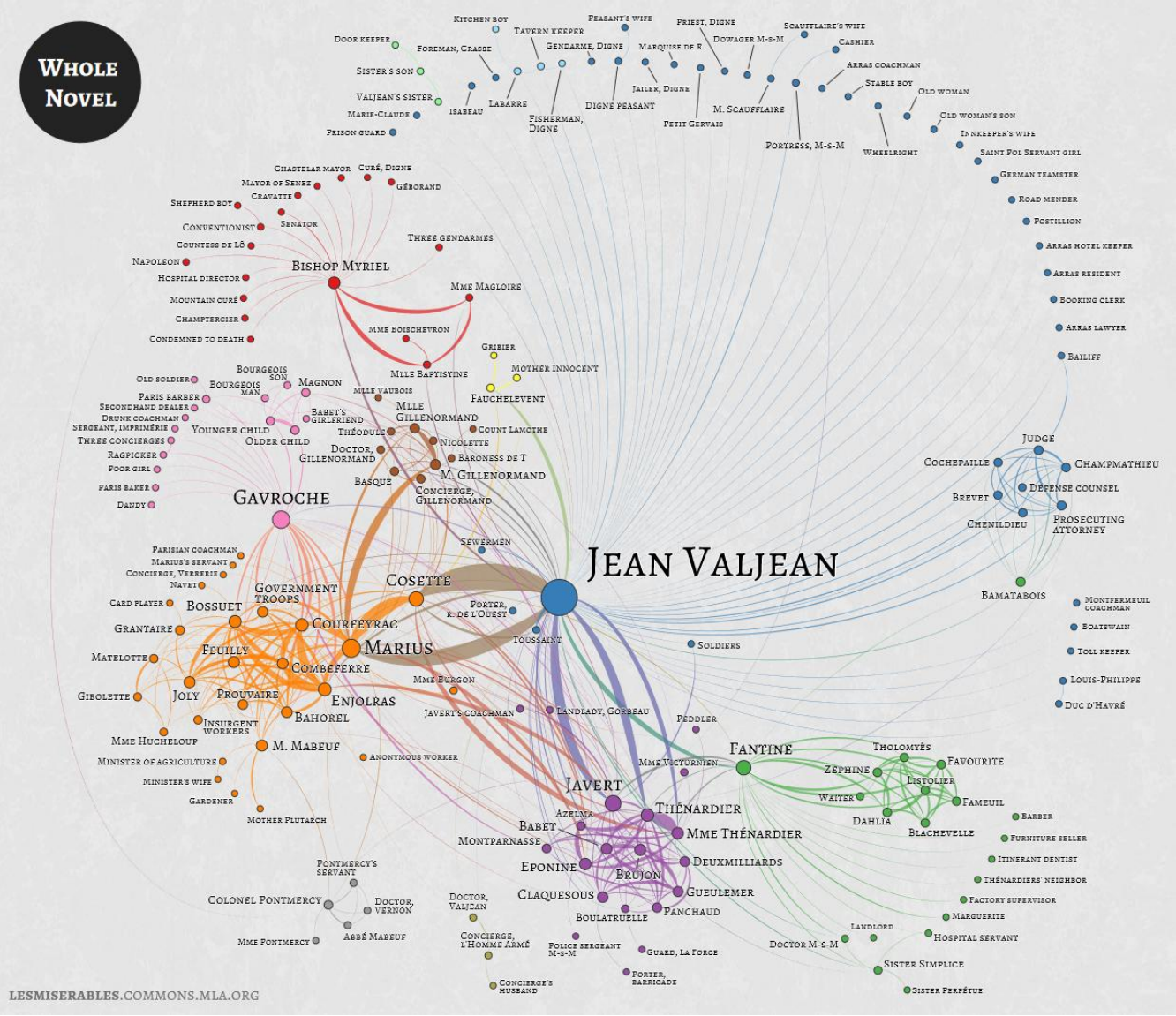
Network Analysis in the Humanities

"In the last few years, literary studies have experienced what we could call the **rise of quantitative evidence.** This had happened before of course, without producing lasting effects, but this time it is probably going to be different, because this time we have digital databases and automated data retrieval […]

What about plot—how can that be quantified? This paper is the beginning of an answer, and the beginning of the beginning is **network theory. This is a theory that studies connections within large groups of objects: the objects can be just about anything—banks, neurons, film actors, research papers, friends . . .**"

Franco Moretti, **"Network Theory, Plot Analysis"** New Left Review, II, 68: 80–102, 2011, p. 81.

Network Analysis in the Humanities

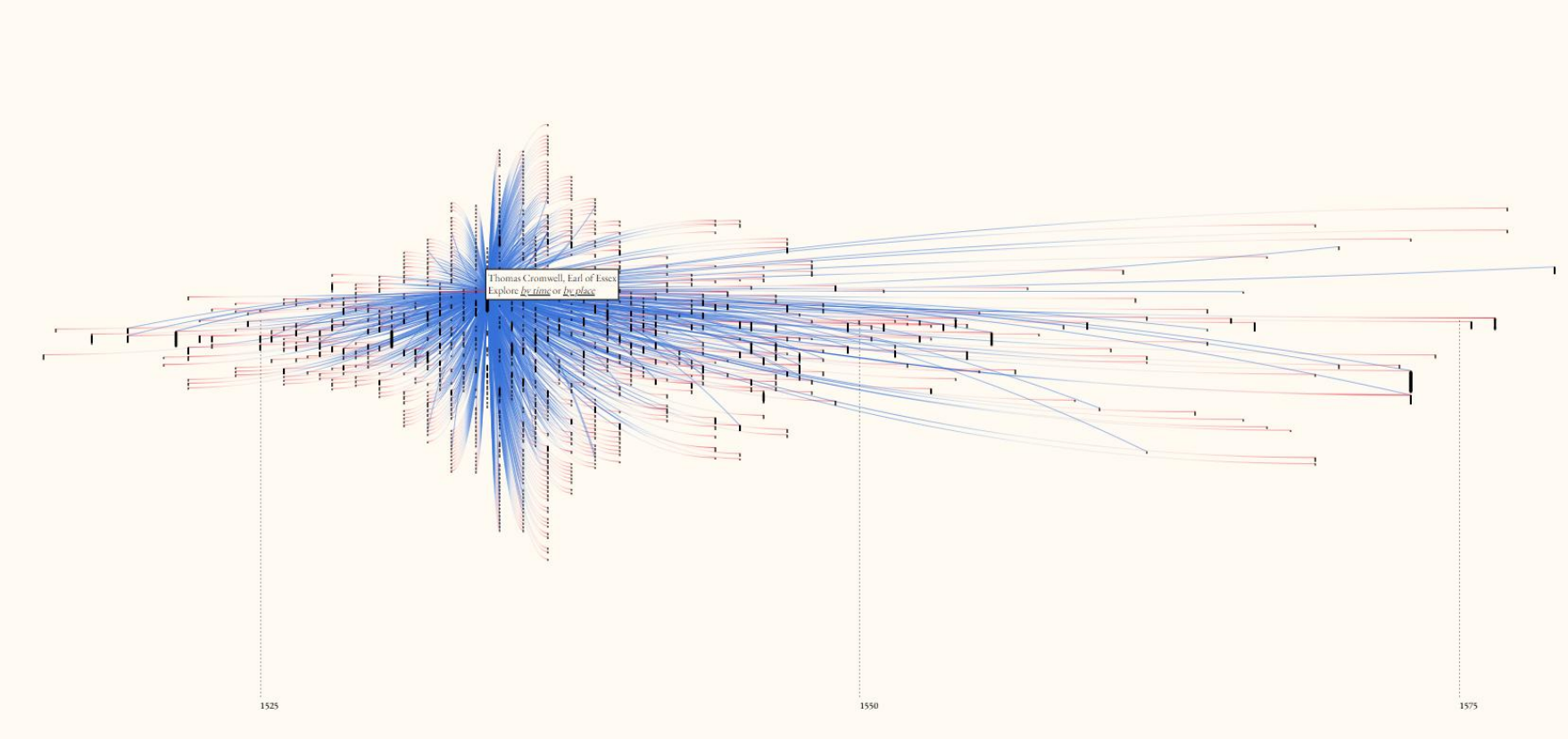


Network visualization of relationships between characters in Les Misérables (1862) by Victor Hugo.

«Visualizing *Les Misérables*», Michal P. Ginsburg

Michal P. Ginsburg, Bradley Stephens, *Approaches to Teaching Hugo's Les Misérables*, Modern Language Association of America, 2018.

Historical Network Analysis

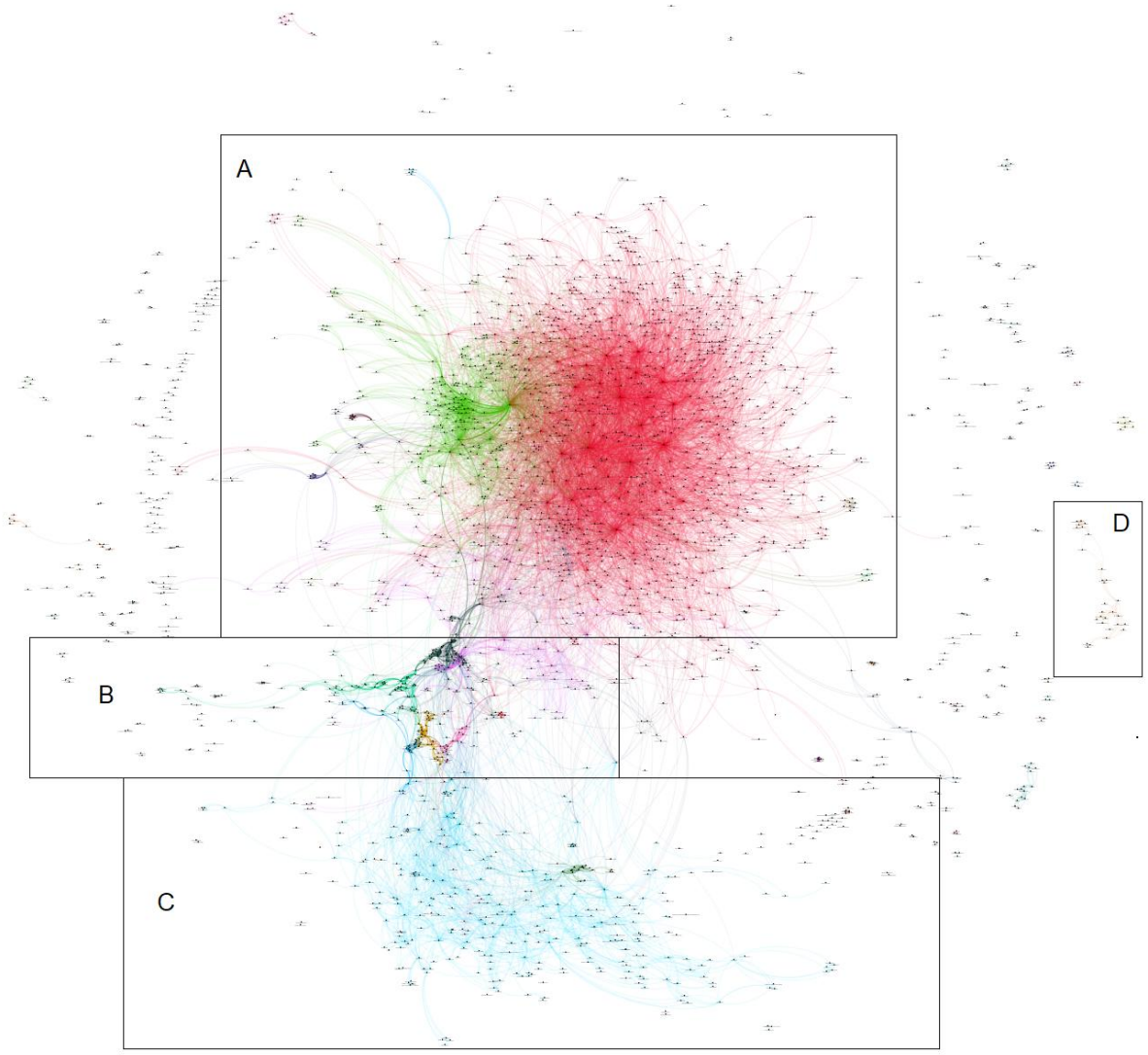


Interactive visualization representing 123.850 letters connecting 20.424 people from the United Kingdom's State Papers archive (1509-1603).

Ruth Anhert, Sebastian E. Anhert, *Tudors Networks of Power*, Oxford University Press, 2023.

Visualization by Kim Albrecht

Historical Network Analysis



Network visualization of relationships between performers and composers in the city of Milan from 1958 to 1962 (examined across 8289 live music performances)

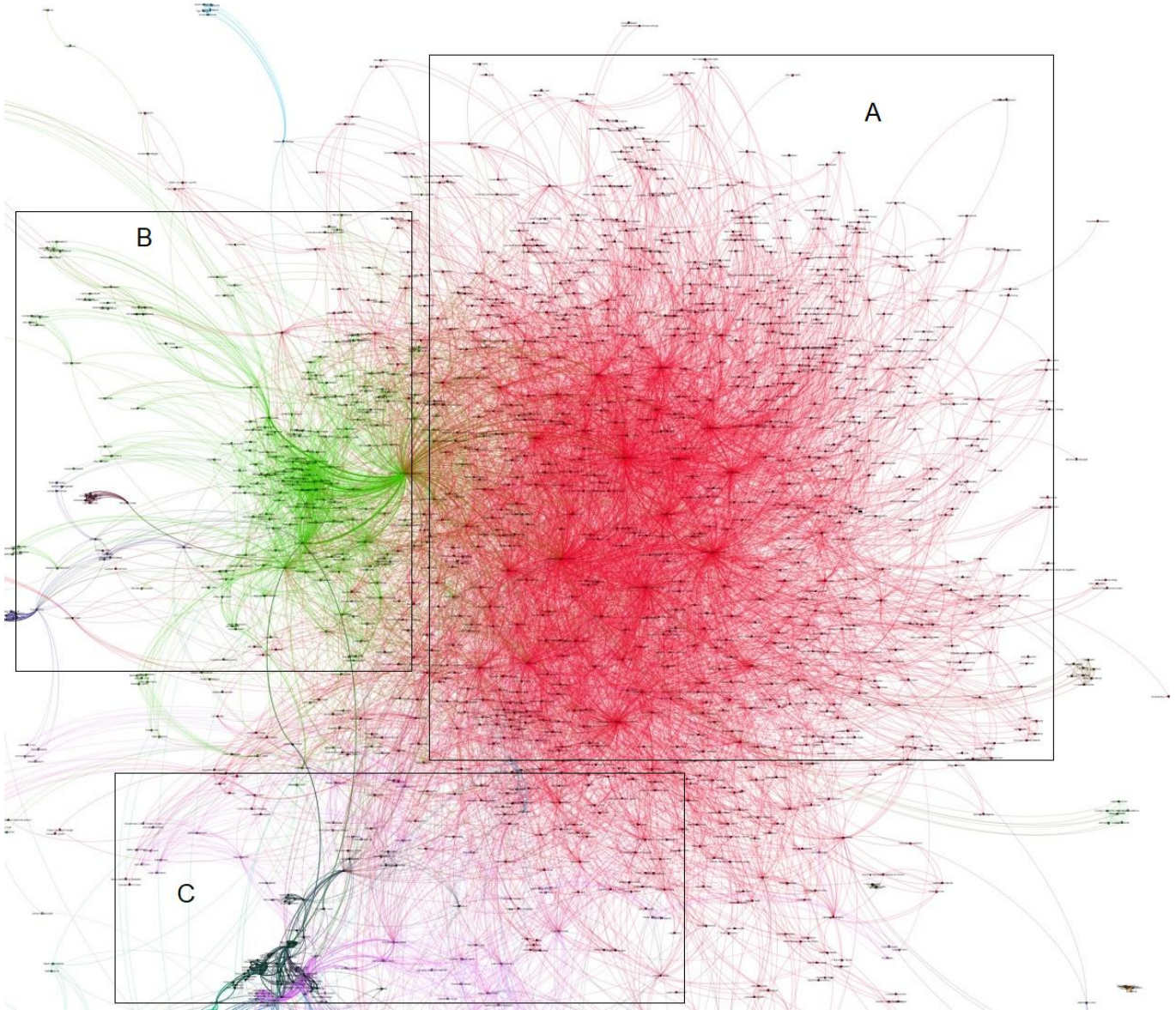
2465 nodes (performers and composers), **11903 weighted connections** (for a total of **52282** links)

Made using Gephi and «Force Atlas 2» layout, which is a force based algorithm capable of considering edge weight

Topology reflects music genre distinctions: classical music (cluster A); scenic music (B); popular music (C)

Martin Nicastro, Milan 1958-1962: Music Topography of a City, 2024

«classical» cluster

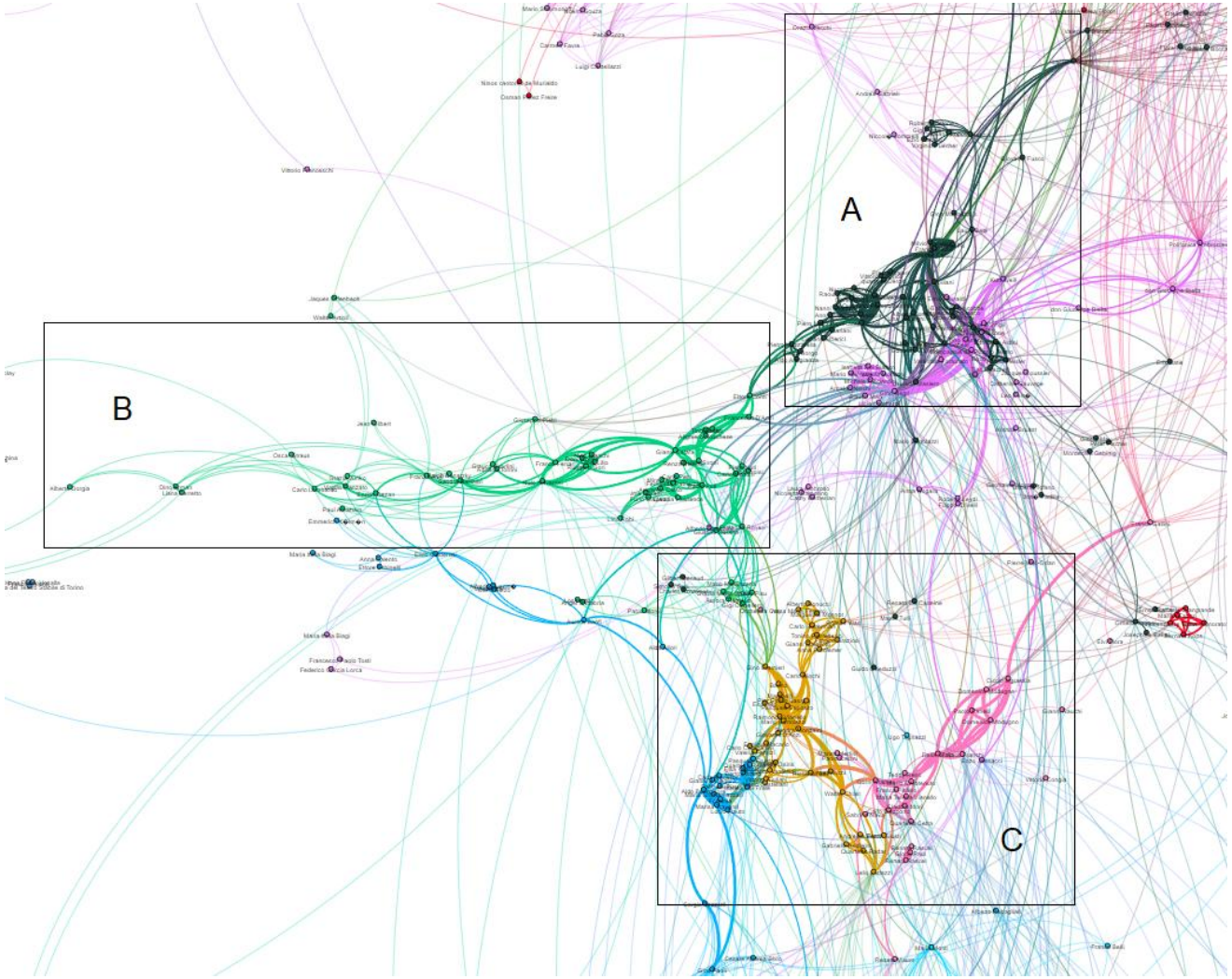


Centre-periphery structure: at the center the composers of the eighteenth-nineteenth century European art music canon

A modularity analysis can identify communities (areas with higher interconnectivity) within a bigger network.

Modularity results:
instrumental music (A, in red); opera (B, in green); early music (C, in purple)

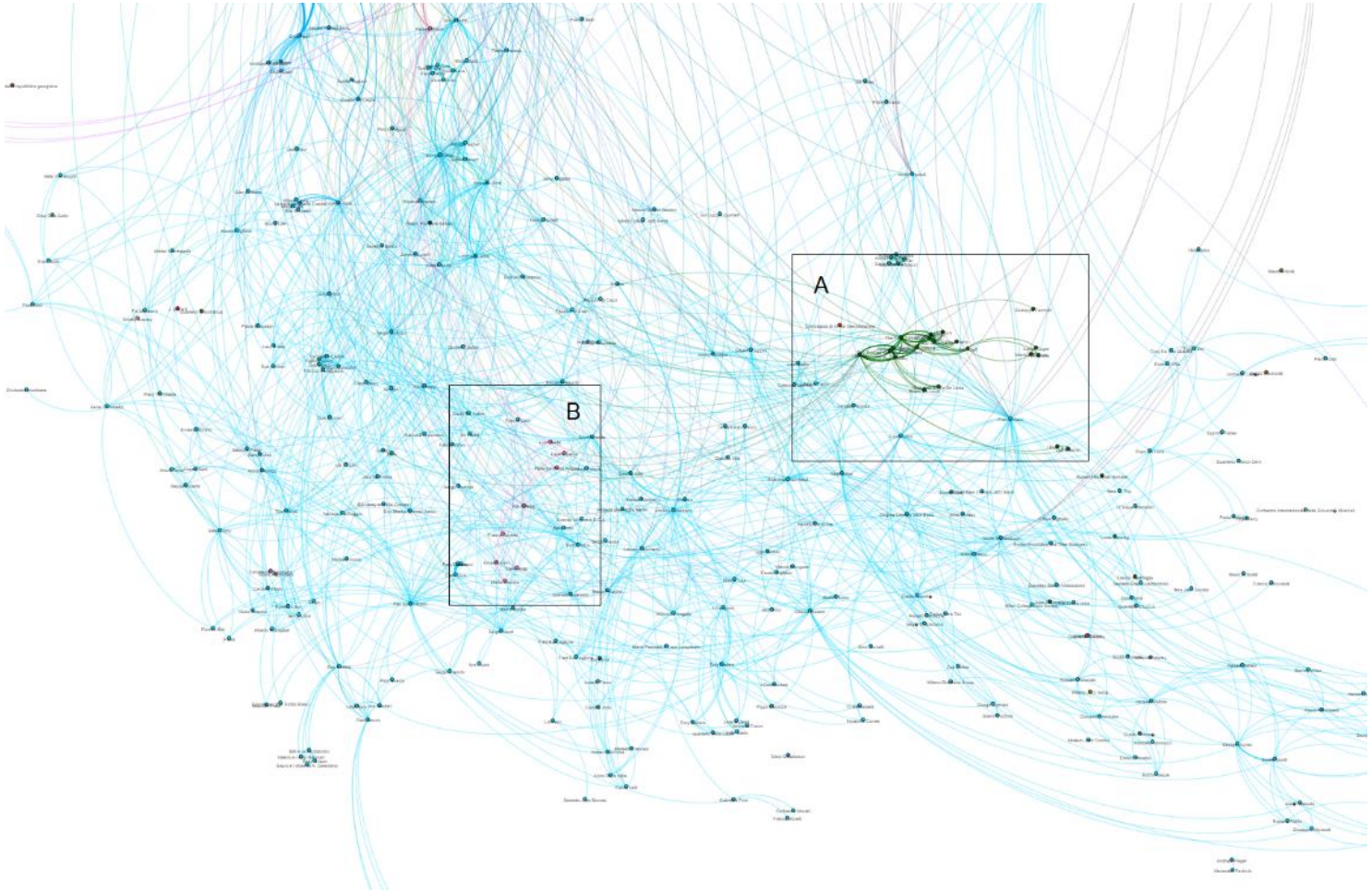
«scenic» cluster



Thread-like and long-limbed network structure, without a core: it acts like a connection portion between the other two clusters

Modularity results: prose theatre (D, petrol), operetta (E, water green), musical comedy (F, ocher and pink)

«jazz/popular» cluster



Compared to the «classical» cluster, it shows a **less dense, more jagged structure, devoid of core**

Modularity results: **no internal division** despite the inclusion of both jazz and popular music practices; this reflects the fact that **at the time jazz musicians were often involved as session musicians for live and studio productions**

Quantitative Analysis

Centrality analyses can detect the asimmetries within a complex network

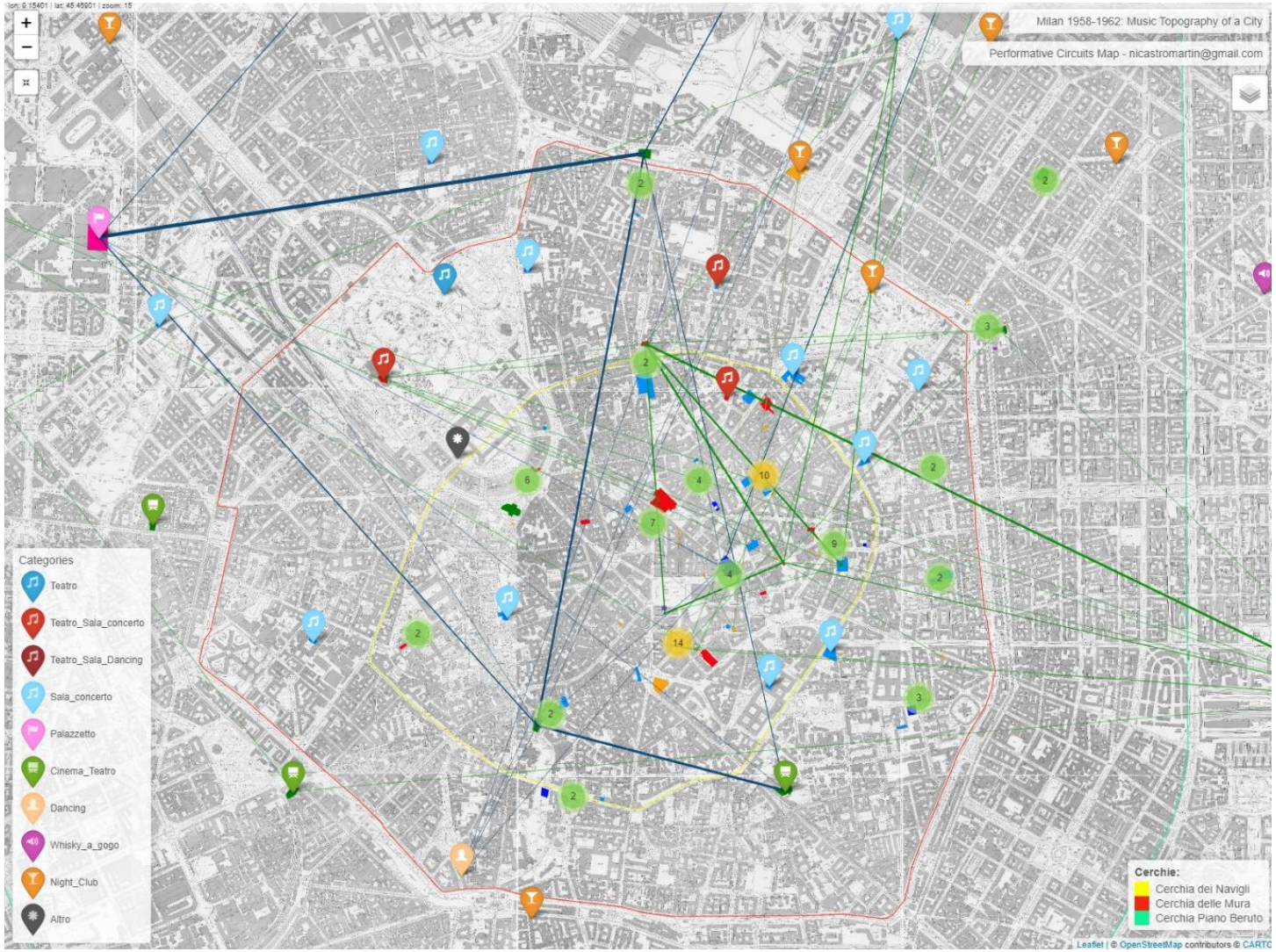
Theatres/concert halls are predominant among venues with **high betweenness centrality values**

Betweenness centrality measures **how many times a node lies on the shortest route between two distant nodes.**

It's often associated with gatekeeping roles (trade routes, political networks) or, more generally, **with having great influence within a network**

Name	**Category**	**Betweenness Centrality**
Teatro Nuovo	Theatre/concert hall	738557.44
Teatro Lirico	Theatre/concert hall	381289.80
Conservatorio G. Verdi (sala Verdi)	Concert hall	346057.04
Teatro alla Scala	Theatre/concert hall	318953.43
Angelicum	Concert hall	282932.51
Teatro Gerolamo	Theatre/concert hall	217562.26
Teatro Manzoni	Theatre/concert hall	137195.64
Supercinema Alcione	Cinema teatro	129045.93
Teatro Olimpia (Olimpia Music Hall)	Theatre/conc. hall/dancing	110411.95
Conservatorio G. Verdi (sala Puccini)	Concert hall	104988.13

Network Analysis and Geospatial Visualization (GIS)



From the hybridization of the two techniques emerges the **role of space**, both urban and performative, **as a relational catalyst**: the concert "theatres/halls" located in the center of the city are the spaces that contribute the most to the formation of the network.

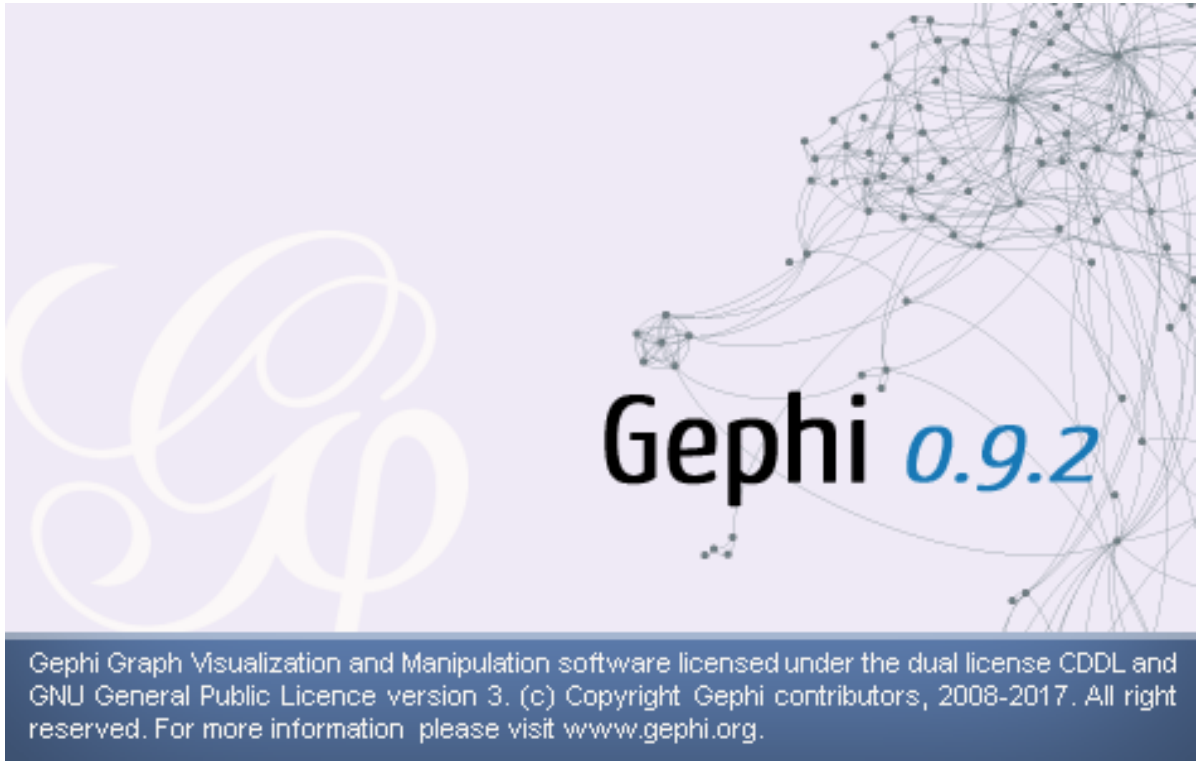
The center of the network corresponds to the center of the city: asymmetries within the network are related to spatial features

Interactive map

How to build a “network graph”

- **Gephi is a free and open-source software,** programmed in JavaScript language, which allows you to visualize and analyze network relationships

- Given its characteristics, it’s one of the reference tools for network analysis in the field of Digital Humanities



File Formats

- **Input** → Data sets in tabular format, the most effective is .csv (Comma Separated Values)

- **Output** → Different types of image files (png, pdf, svg), but also georeferenced files that can be imported into software dedicated to digital mapping (kml)

Inputs

Nodes

	A	B	C	D	E	F
1	ID	Label	Attributo	Generi	Latitude	Longitude
2	033p	Aida Accolla	1	Balletto	45.4642035	9.189982
3	1024p	Liliana Cosi	1	Balletto	45.4642035	9.189982
4	1059p	Luciana Savignar	1	Balletto	45.4642035	9.189982
5	068p	Alfredo Adami	2	Balletto, Canzone internazi	41.9027835	12.4963655
6	1882p	Xavier Cugat	5	Balletto, Canzone italiana,	41.9794005	2.8214264
7	849p	I Brutos	37	Balletto, Canzone italiana,	45.0703393	7.686864
8	1662p	Roy Bradley	31	Balletto, Canzone italiana,	51.5476301	-0.2369246

Edges

	A	B	C	D
1	Source	Target	Type	Weight
2	001p	028p	Undirected	1
3	001p	1386p	Undirected	1
4	001p	1788p	Undirected	1
5	002c	062c	Undirected	1
6	002c	389c	Undirected	1
7	002p	512c	Undirected	6
8	002p	530c	Undirected	6

The columns highlighted in red are mandatory. Changing the names or missing columns will cause major errors during the import process.

Optional columns can be additional attributes linked to both nodes (categorical or quantitative) and edges.

The use of coordinates as node attributes allows networks to be geolocalised.

Datasets mandatory columns

Nodes:

1. **ID** is an unique identifier (it can be textual, numerical or alphanumerical) for each node.

2. **Label** is the text that Gephi will attach to each node.

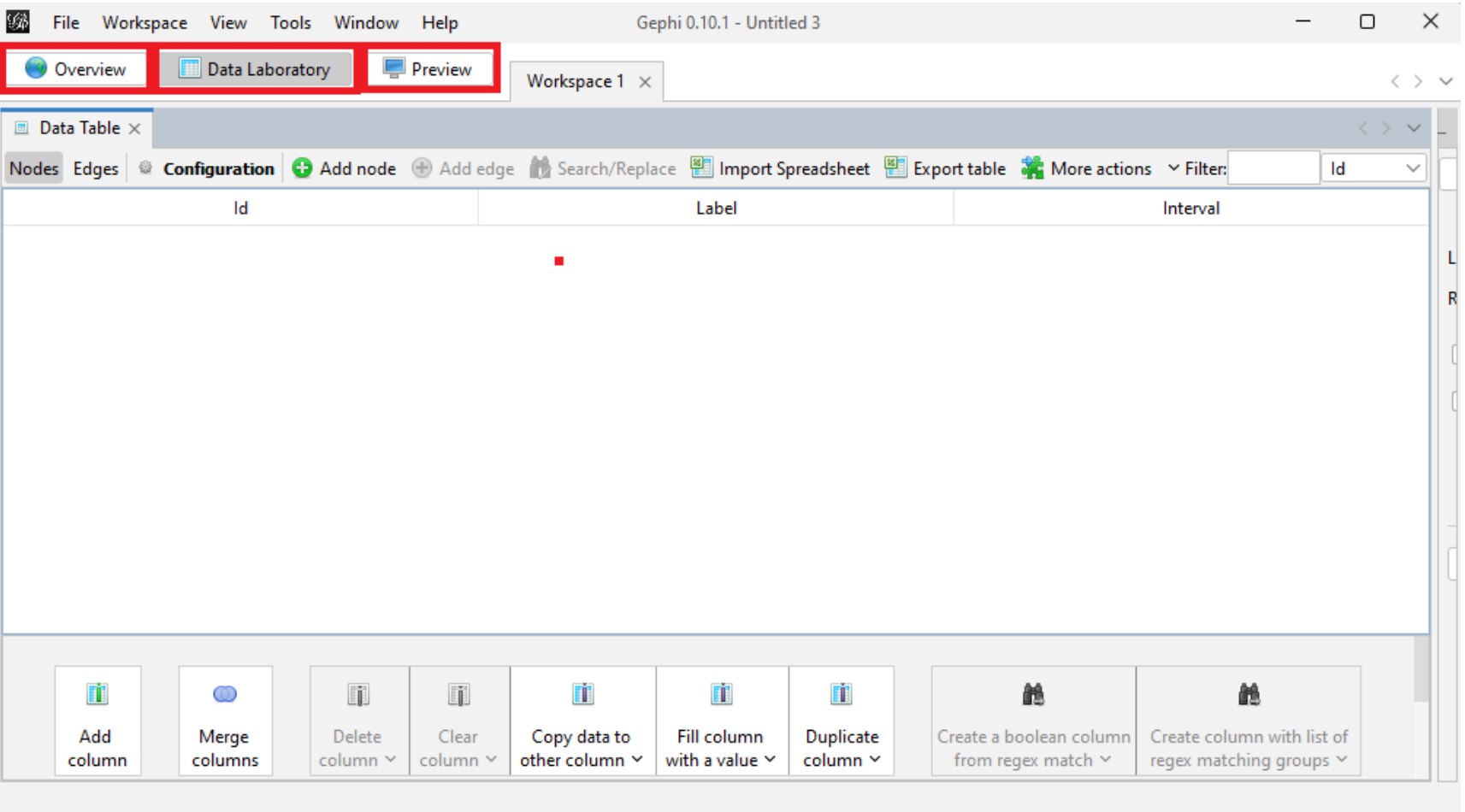
Edges:

1. **Source** is the first node involved in the tie.

2. **Target** is the second node involved in the tie. Important: **both Source and Target values must correspond to the ID values entered in the nodes table.**

3. **Type** only accepts two values: “**Directed**” (univocal relation) or “**Undirected**” (biunivocal relation)

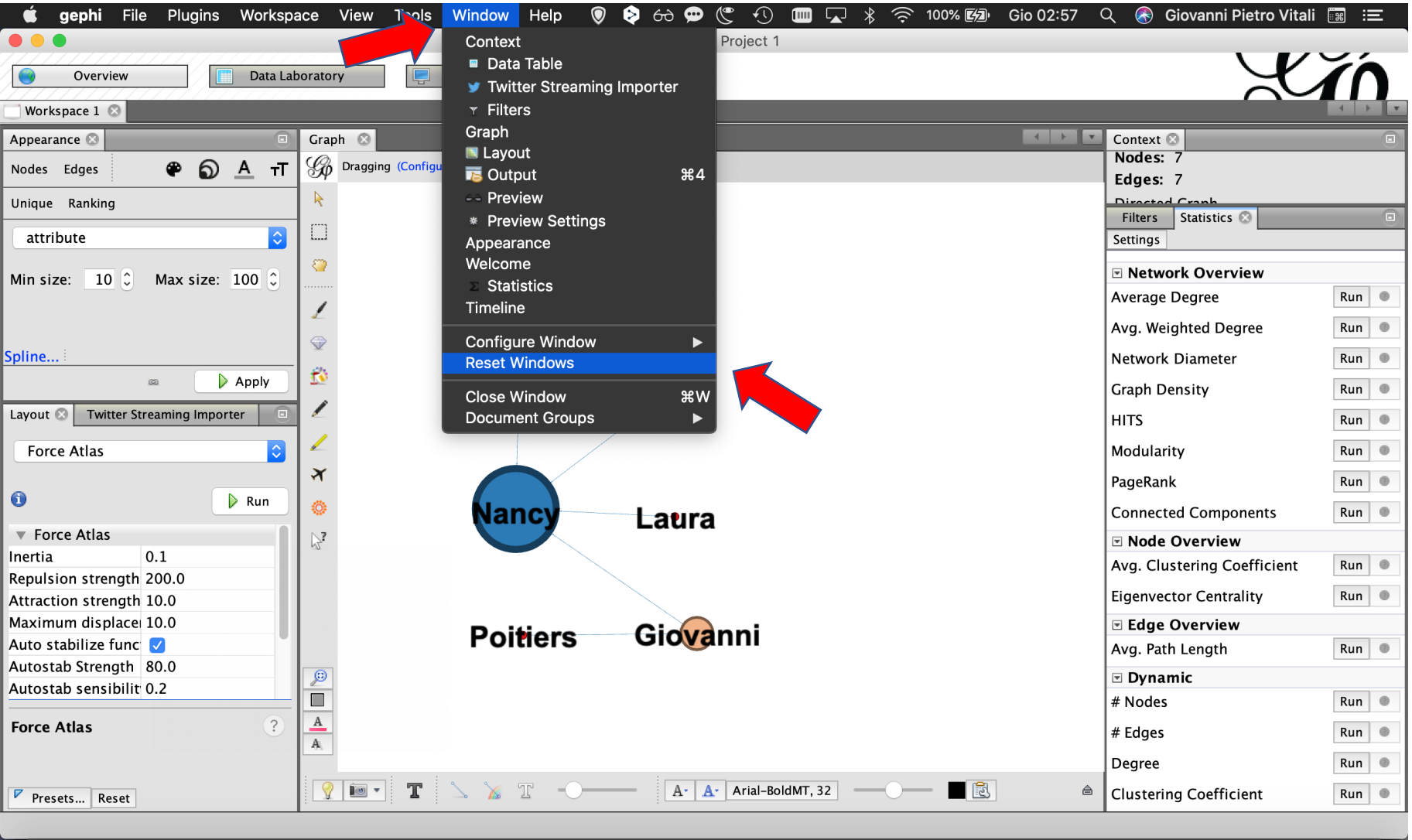
The workspace



- **Data laboratory** contains the imported tables
- **Overview** is where we run analysis and spatialization algorithms; it also allows a first phase of graphic customization (size, colours)
- **Preview** shows the graph as it will be exported. It's where the final graphical touches are made

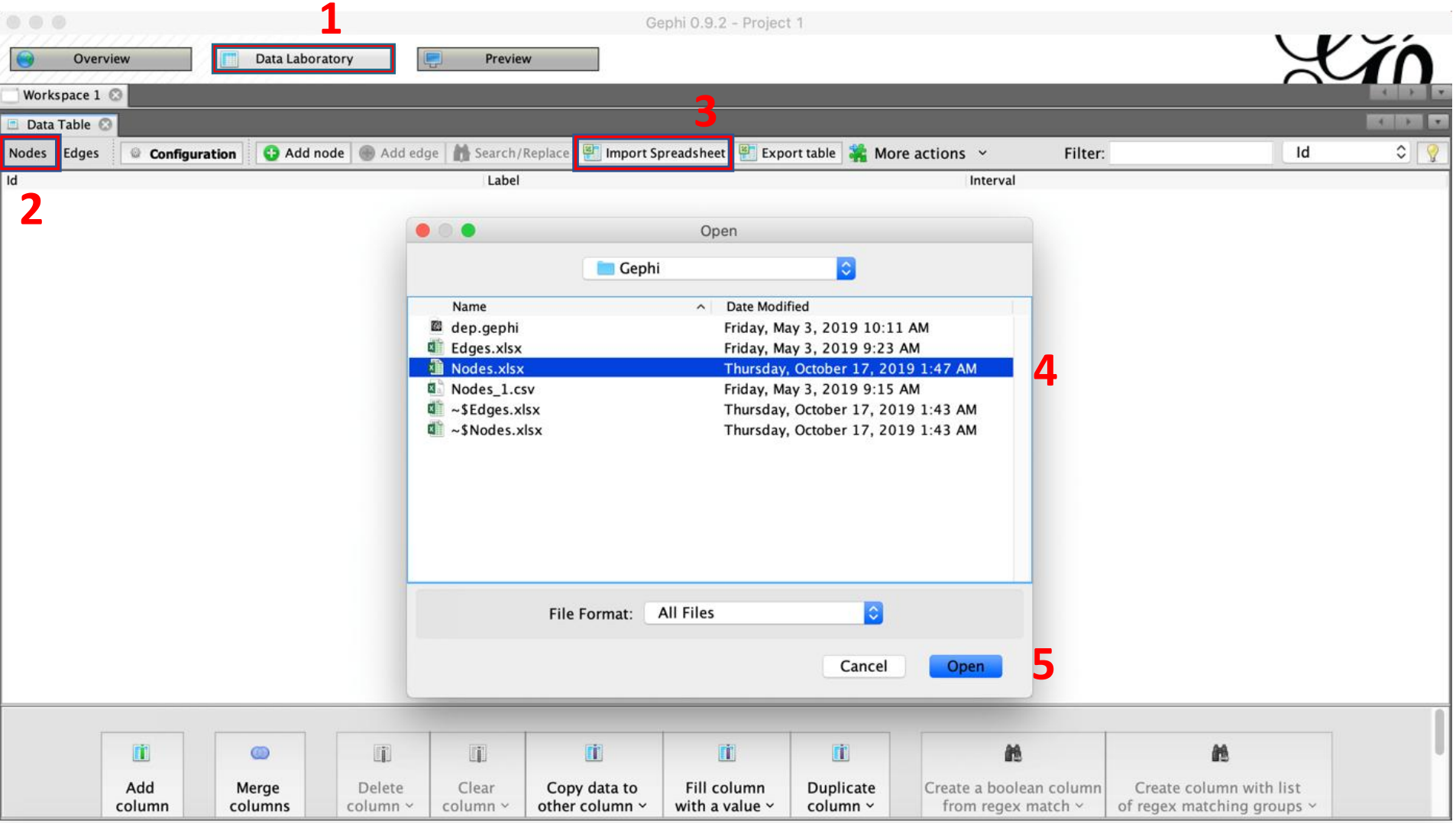
An effective use of the software consists in moving continuously between the three windows

Reset windows on Mac



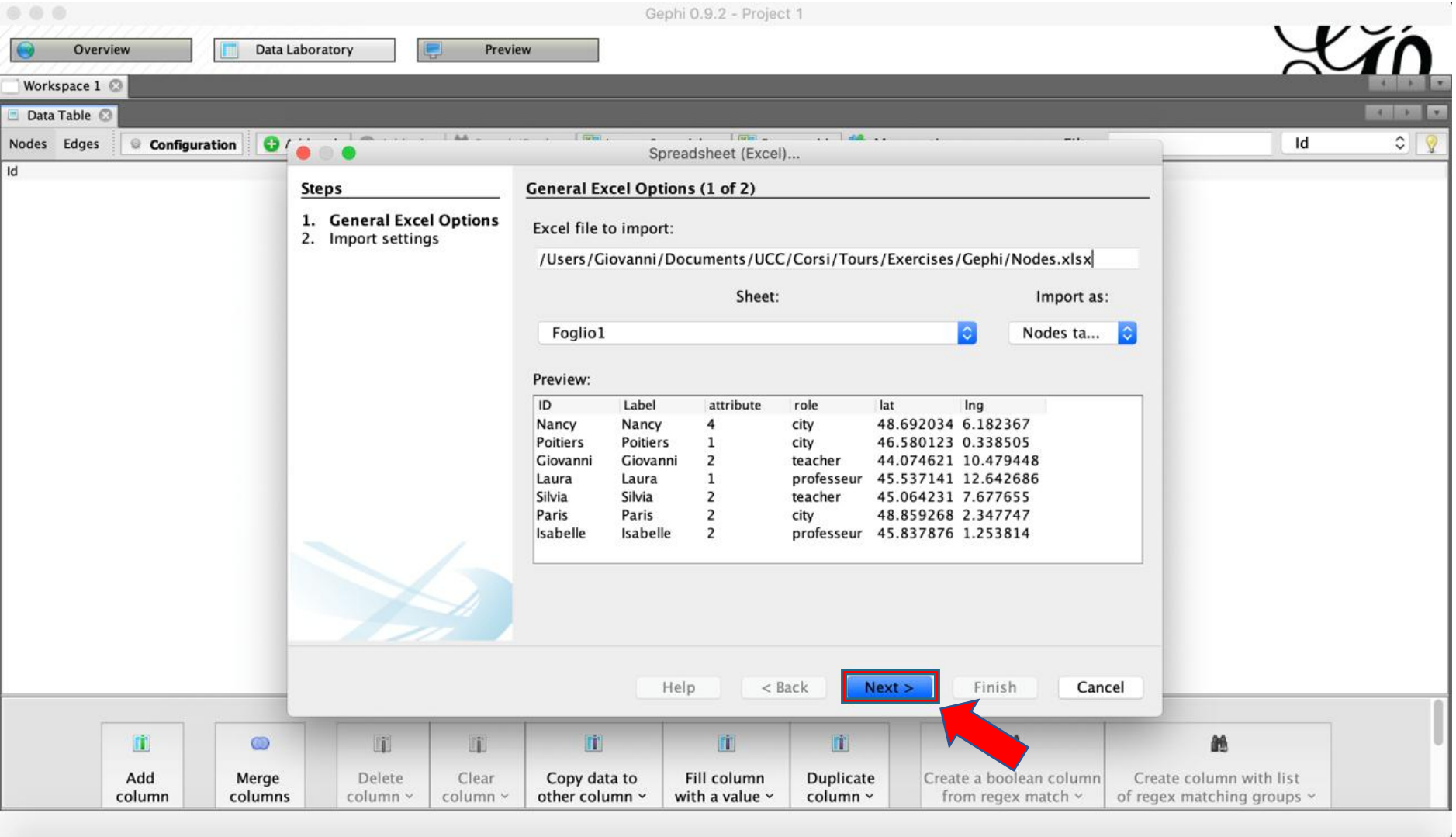
- In the Mac version of the software windows are not docked. Don't worry: there is a specific command to reset them!

Data Laboratory: importing nodes

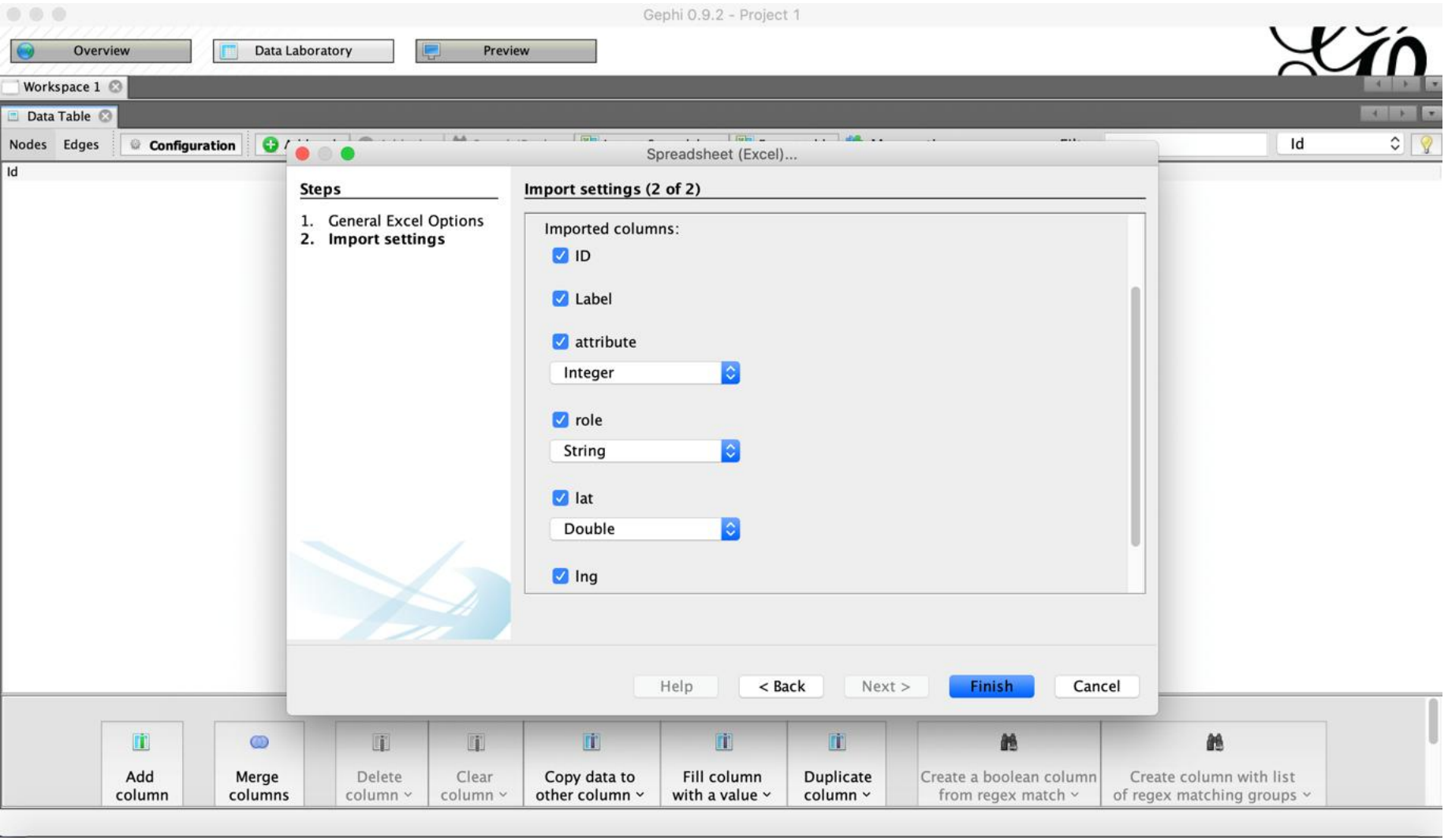


- The starting point of each project is always the **Data Laboratory** window where you can import your data. **It is good practice to start with the node table**.

Preview of imported data (1)



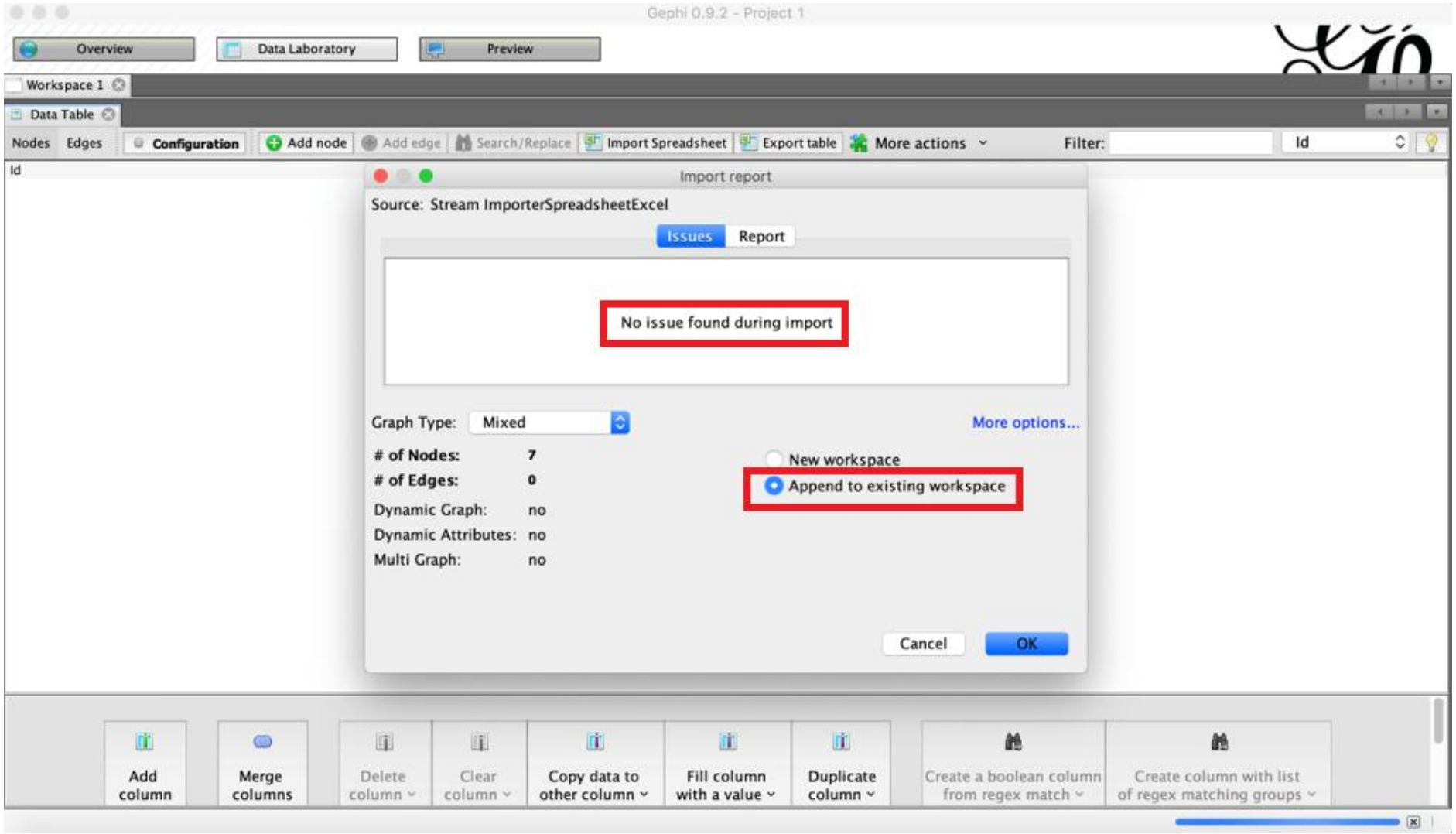
Preview of imported data (2)



- **Gephi automatically identifies the format of each column:** here you can check that they are correct. Note how the «ID» and «Label» columns cannot be modified

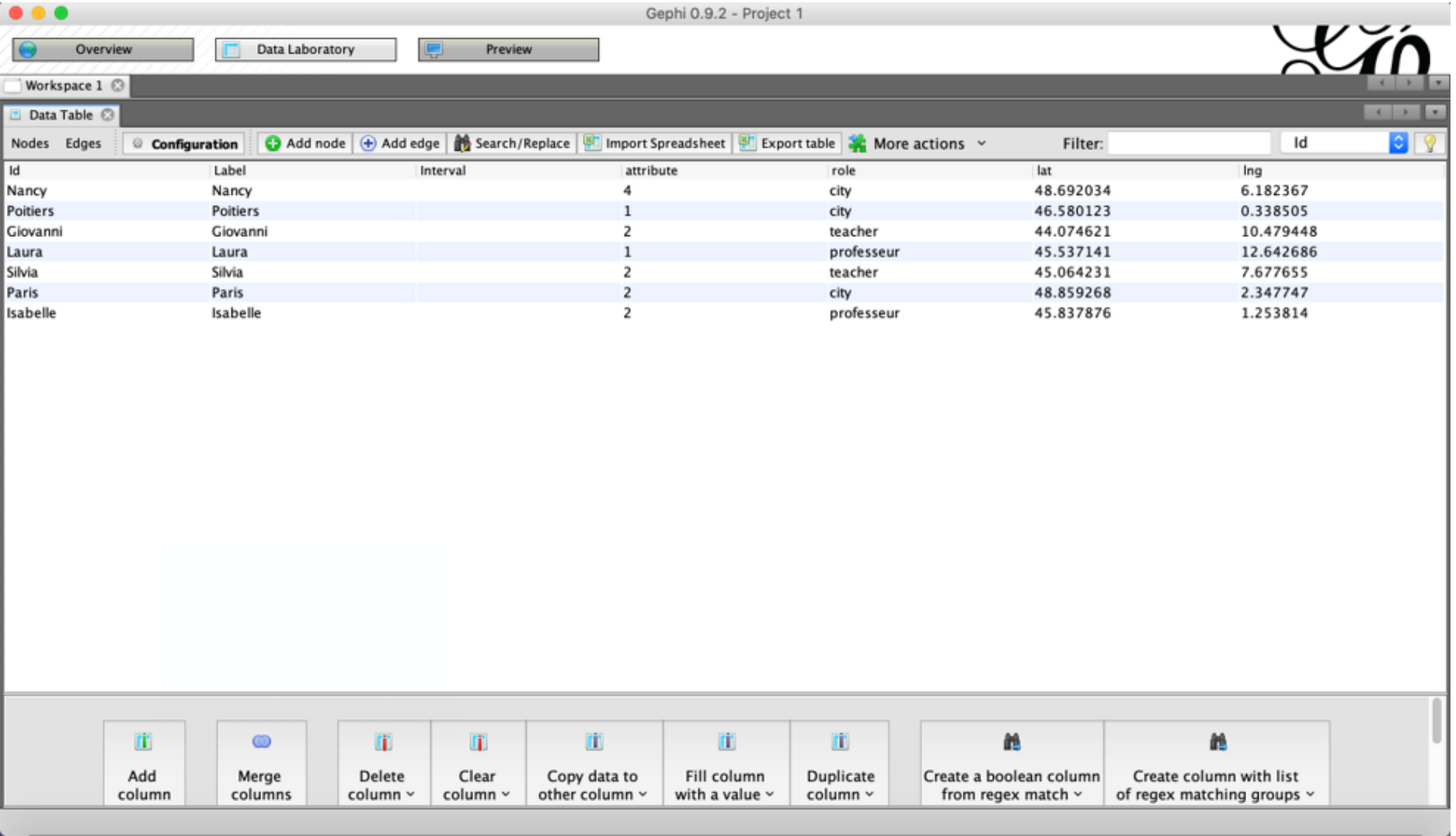
- Latitude and Longitude must always be in the “double” format

Import report



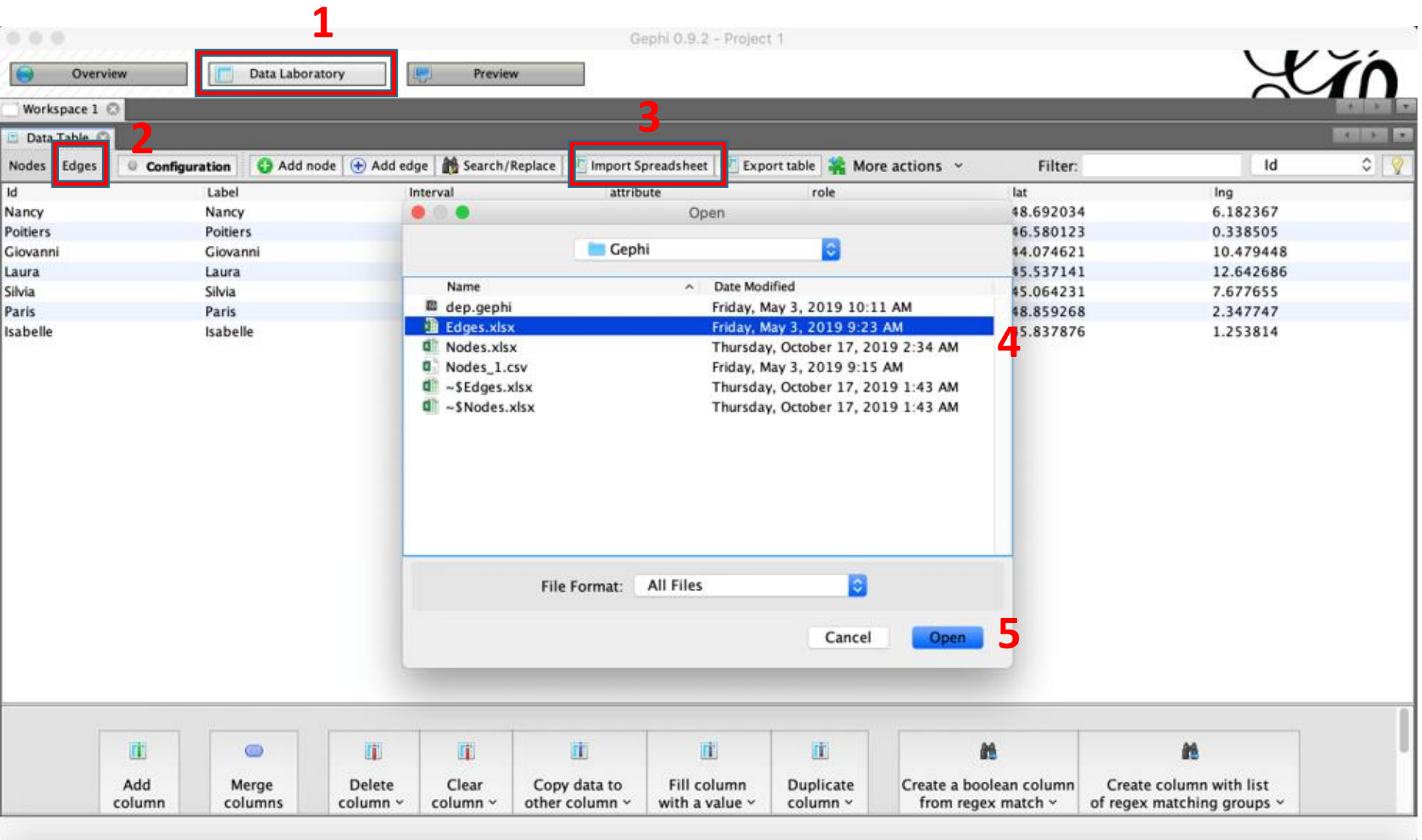
- **Attention!** In the current version the default choice is to import each data set into a different workspace: **it is important to select the "append to existing workspace" option**

Tabular view within the Data Laboratory



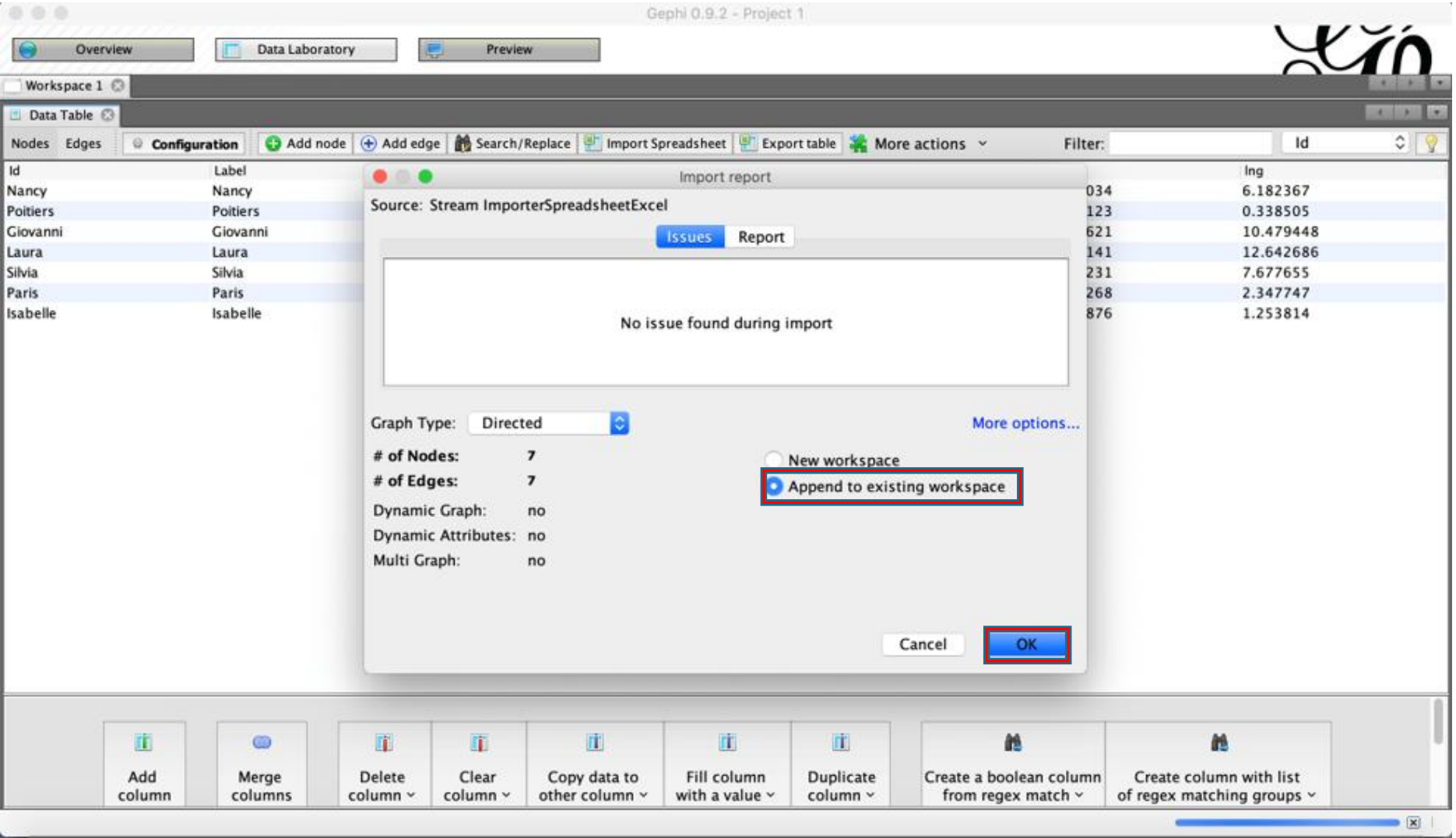
- At the end of the import process (if carried out correctly) you can consult the table in the Data Laboratory window

Data Laboratory: importing edges



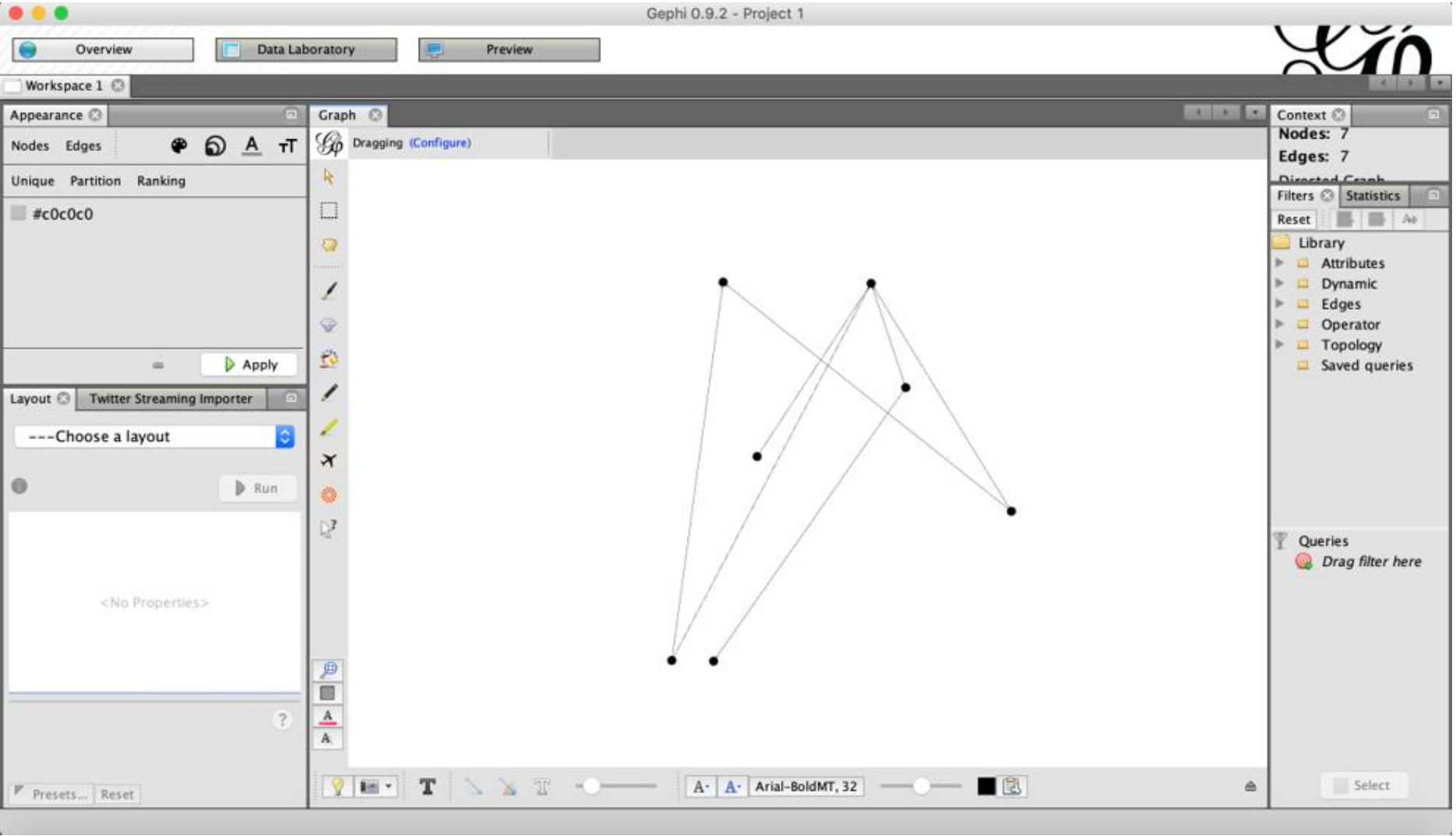
- Repeat the same steps with the edges table

Import report: edges



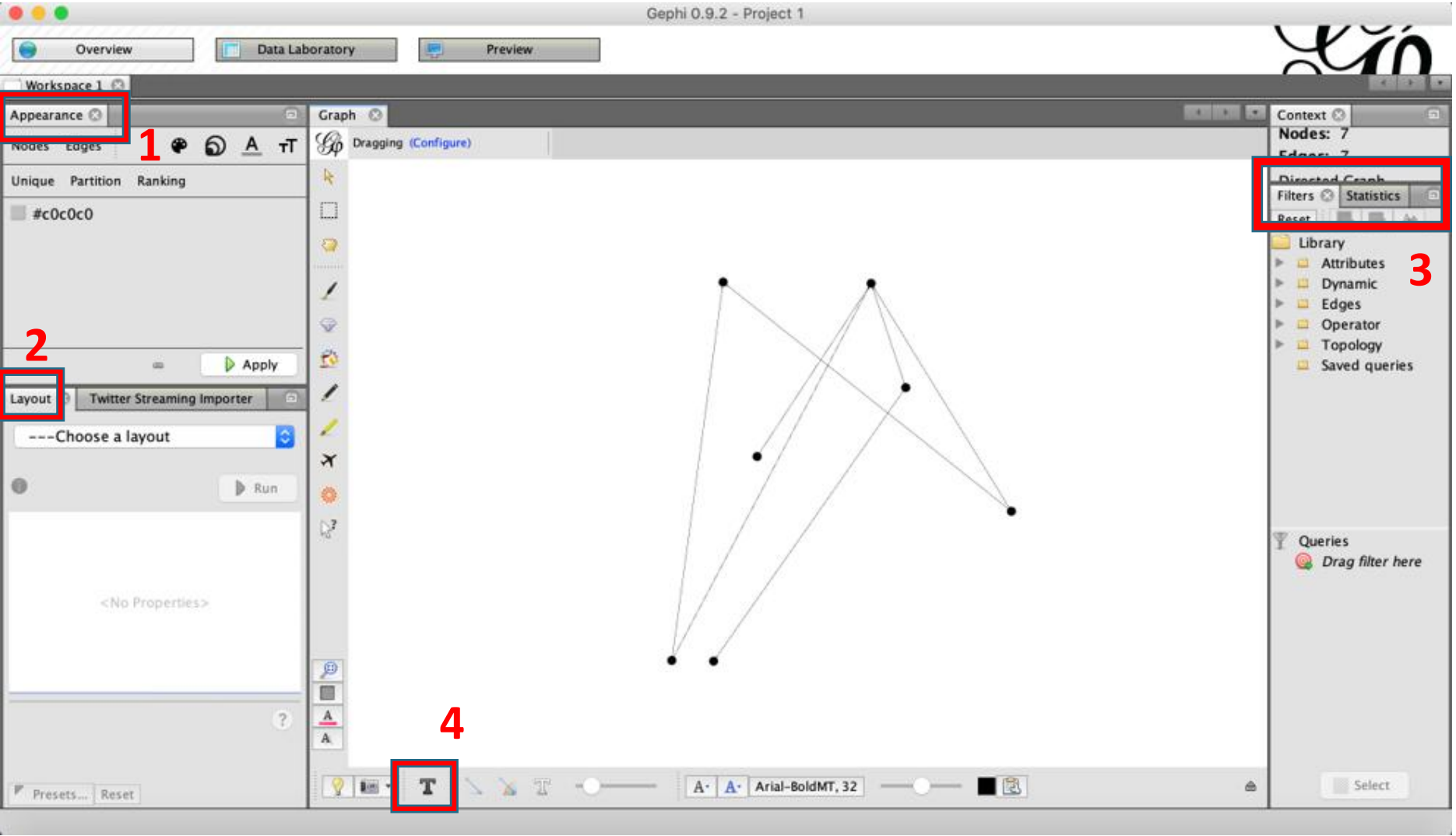
- **Attention!** As before, it is important to select the **«append to existing workspace»** option

Overview



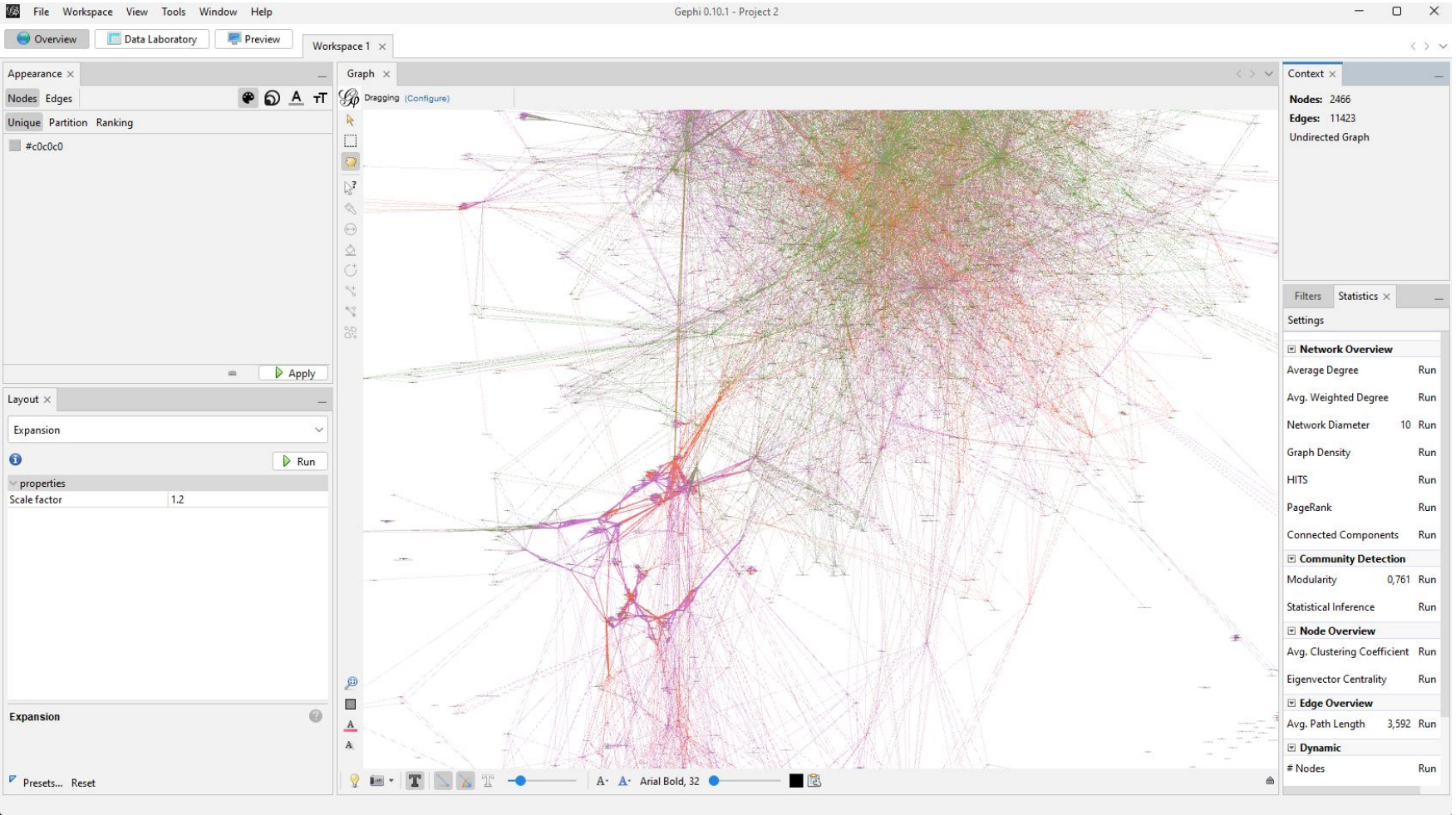
- **Let's move to the Overview window**: if the import of the data sets was carried out correctly you will see a graph with a random spatial layout.

Overview selectors



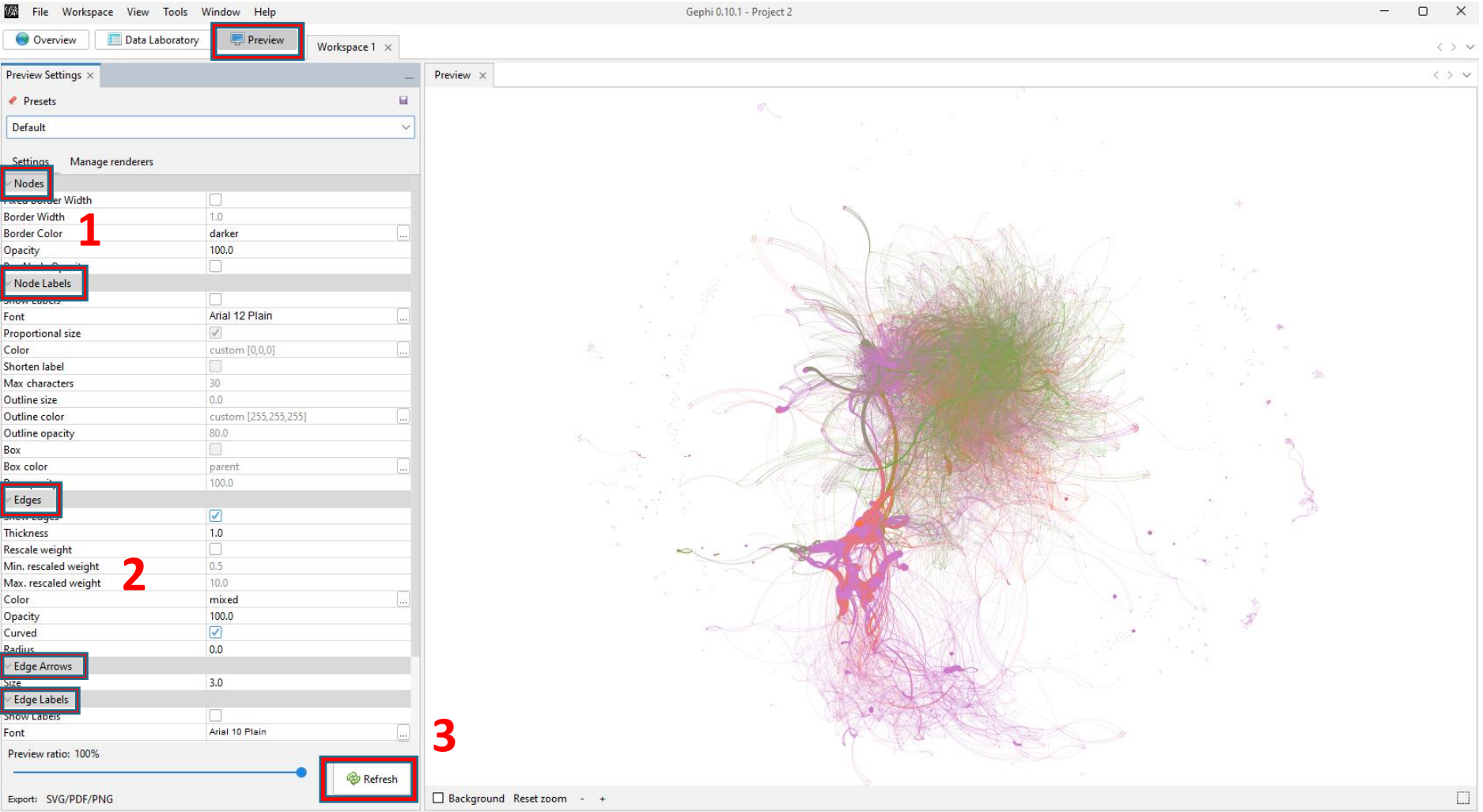
1. **Appearance** (graphic costumization for nodes and edges)
2. **Layout** (spazialization algorithms)
3. **Filter/Statistics** (Analysis algorithms and filters)
4. Click here to activate labels!

Customize your graph



- Once you have customized the graphical appearance of the graph (**spatial layout, colors and dimensions, filters**) check the results in the Preview window

Preview window

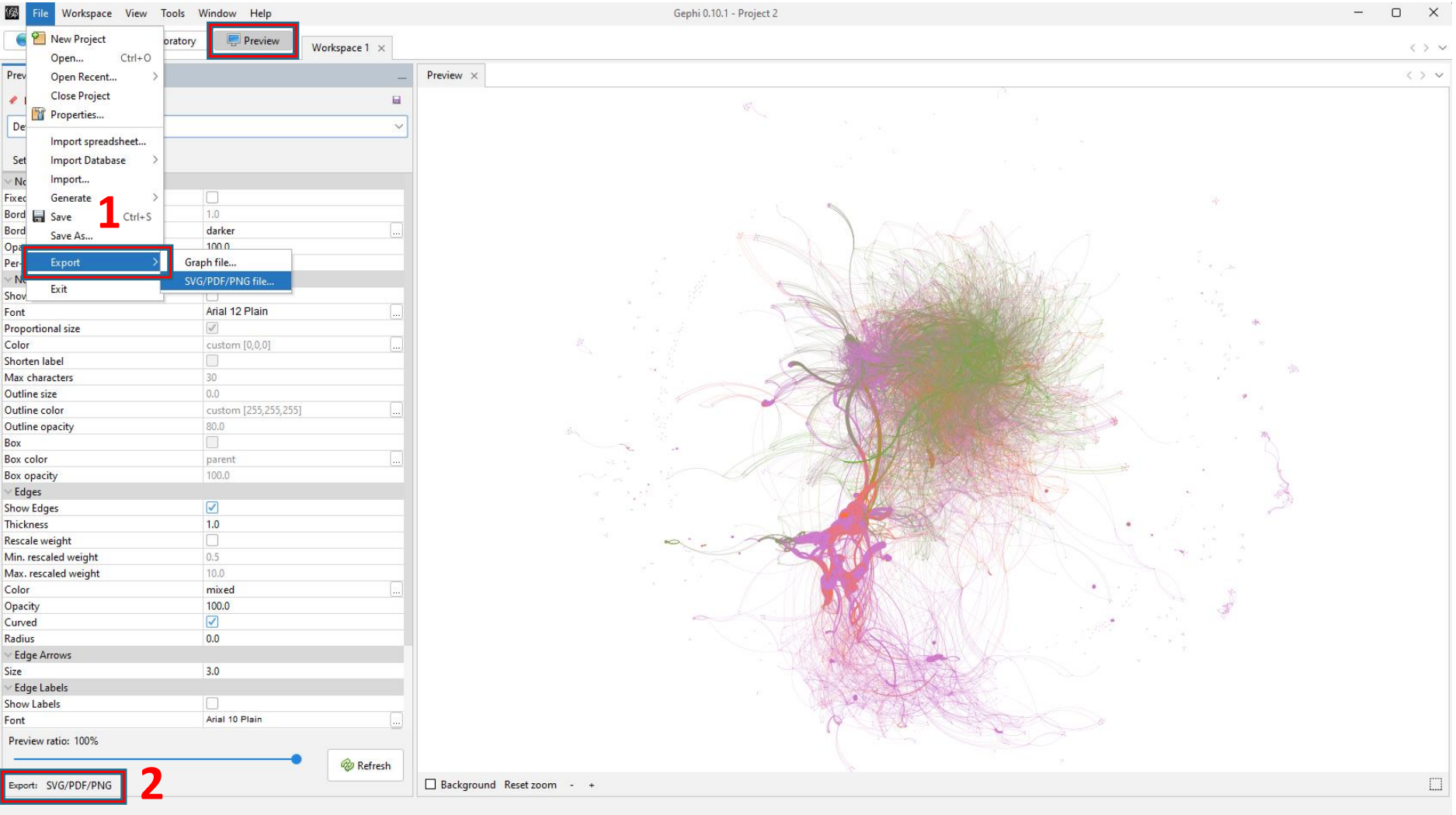


1. Nodes final graphic customization (**labels, proportional size**)

2. Edges final touches (**color, curved vs straight, arrows, thickness, labels**)

3. Important: **remember to press the «refresh» button (3)** to update the visualization!

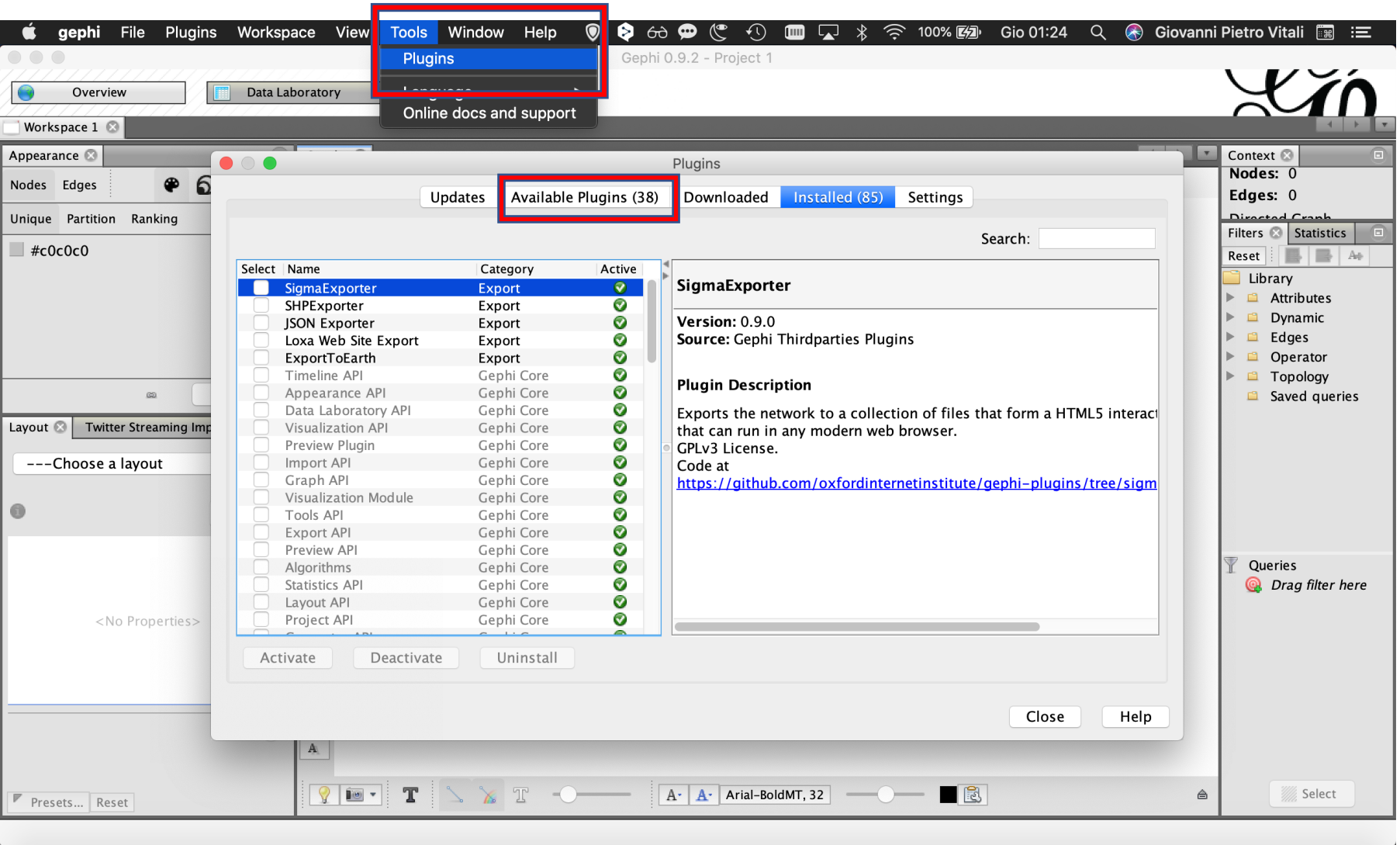
Export



1. **First option: from the file menu.** It can be also used to export the graph in a geographical format («graph file-kmz»)

2. **Second option: direct export.** Only for image files!

Plugins



Select and install:

1. **Geolayout**
2. **Export to Earth**
3. **MultiGravity Force Atlas 2**