



GEPHI TUTORIEL

<https://gephi.org/>

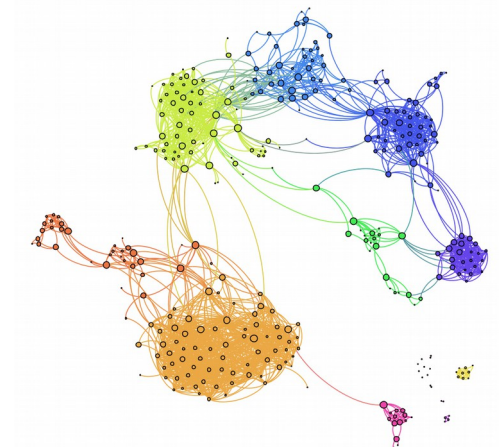
<https://github.com/gephi/gephi>

Gephi est un logiciel libre d'analyse et de visualisation de réseaux, développé en Java et fondé sur la plateforme NetBeans.

L'un des principaux intérêts de l'utilisation de Gephi pour cartographier des données est la possibilité d'utiliser de nombreux calculs liés à la théorie des graphes pour les appliquer aux données utilisées. On peut ainsi visualiser quels sont les éléments d'un réseau les plus centraux, les plus éloignés, les mieux connectés, etc. De nombreux plugins existent également pour permettre de cartographier des ensembles de données très variés.



Liens entre membres d'une organisation



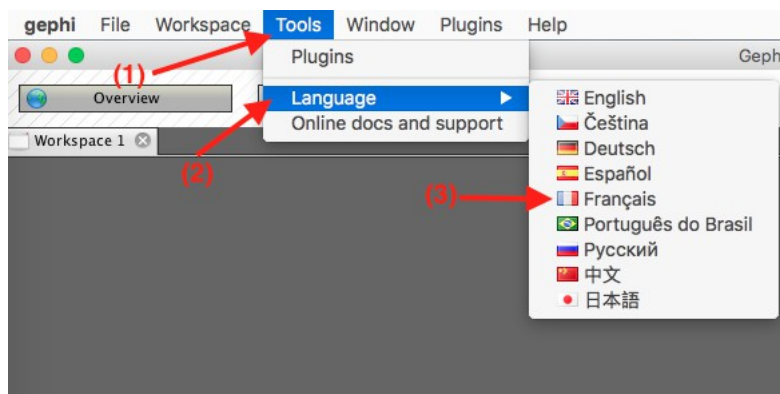
Réseau social d'un utilisateur de Facebook.

internationale

Préparation de l'espace de travail

Changer la langue

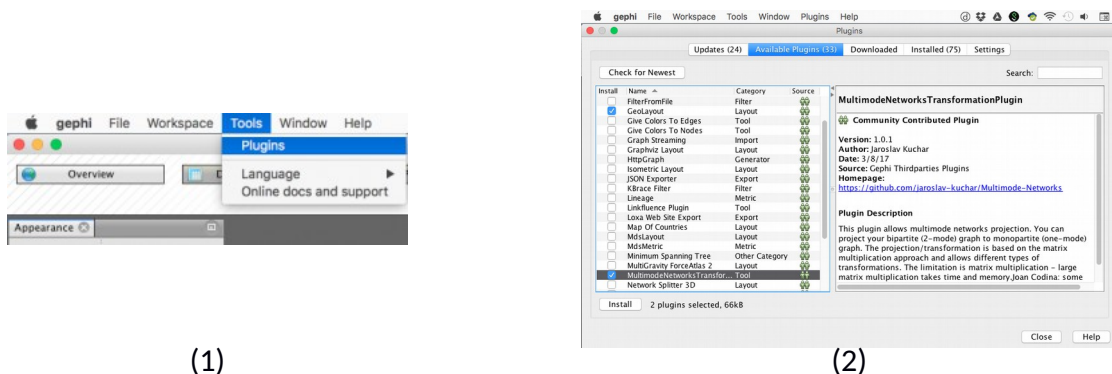
Vous pouvez changer la langue de Gephi en cliquant sur **Tools** (1) > **Language** (2) puis en sélectionnant la langue de votre choix. Pour ce tutoriel nous avons choisi le français (3).



Installer les plugins

Afin d'aller au-delà des fonctionnalités basiques du programme, nous allons installer quelques plugins supplémentaires. Pour aller dans le menu des Plugins, suivre **Tools** puis **Plugins** (1). Dans la fenêtre qui s'ouvr, aller à l'onglet **Available Plugins** (2) et choisir :

- « GeoLayout »
- « SigmaExporter »
- « Loxa Web Site Exporter »
- « JSON Exporter »
- « ExportToEarth »
- « SHPEXporter » et « MultimodeNetworkTransformation »



C'est quoi un réseau ?

Un réseau est constitué de deux composantes : une liste des acteurs composant le réseau et une liste des relations entre ces acteurs. Les premiers seront alors appelés « nœuds » (*nodes*) et les seconds seront appelés « arêtes » (*edges*).

Un jeu de données typique pour travailler sur Gephi est donc normalement composé de deux fichiers : un de *nodes* (1) et un autre d'*edges* (2).¹

	A	B	C	D	E	F
1	id	Label	Attribute1	City	Latitude	Longitude
2		376088951 name1		1 Berlin	52.51	13.4
3		17647430 name2		2 Berlin	52.51	13.4
4		32416061 name3		1 Berlin	52.51	13.4
5		550180187 name4		2 Berlin	52.51	13.4
6		28110685 name5		1 Berlin	52.51	13.4
7		14845783 name7		1 Berlin	52.51	13.4
8		63293 name8		2 Berlin	52.51	13.4
9		22606966 name9		1 Berlin	52.51	13.4
10		193763394 name10		2 Berlin	52.51	13.4
11		191004748 name11		1 Berlin	52.51	13.4
12		61502712 name13		1 Berlin	52.51	13.4
13		2230301 name16		2 Berlin	52.51	13.4
14		17670064 name17		1 Berlin	52.51	13.4

(1)

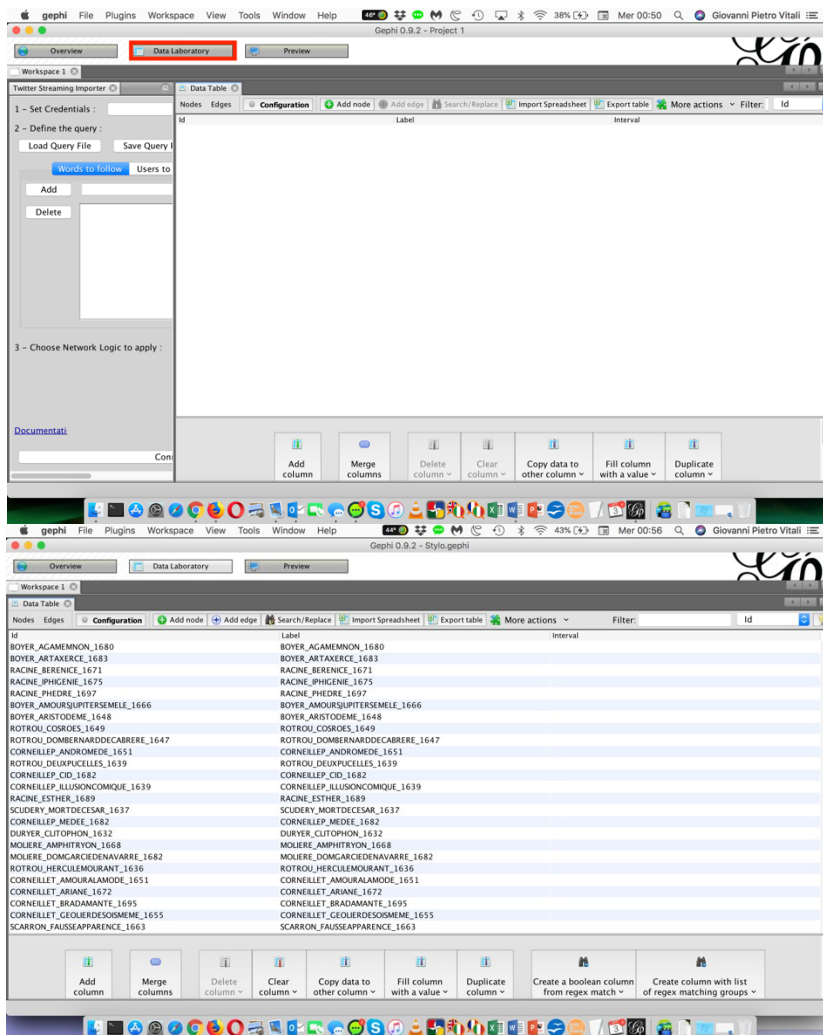
	A	B	C	D
1	Source	Target	Type	Weight
2	376088951	17647430	Directed	1
3	376088951	32416061	Directed	1
4	376088951	550180187	Directed	2
5	376088951	28110685	Directed	3
6	376088951	870137221	Directed	3
7	550180187	376088951	Directed	2

(2)

¹ Notez qu'il est possible de ne travailler qu'avec des fichiers d'*edges* et de générer les *nodes* à partir des liens.

Onglet 1 : *Data Laboratory* (laboratoire des données)

Gephi fonctionne avec trois sections qui sont interconnectées et doivent être utilisées conjointement pour réussir la transformations des données en graph.

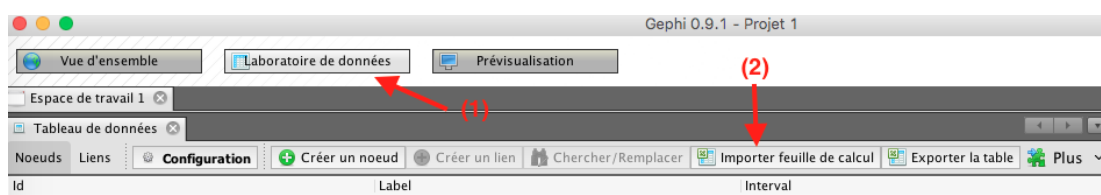


La section *data laboratory* permet d'insérer les données qui seront visualisées graphiquement dans l'*overview*.

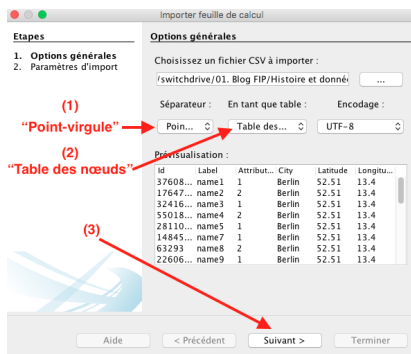
ATTENTION !
Tout ce qui est modifié dans le *data laboratory* est immédiatement modifié graphiquement dans l'*overview*

Importer les données dans Gephi

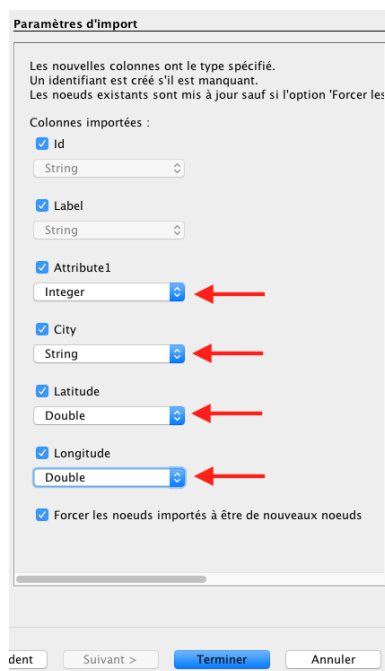
Au démarrage de Gephi, cliquer sur Nouveau projet puis dans Laboratoire de données (1) cliquer sur Importer feuille de calcul (2).



Nodes



Dans Choisissez un fichier CSV à importer, sélectionnez le fichier contenant les nœuds. Dans les options générales d'importation, choisir que la séparation entre les colonnes soit exprimée en point-virgule (1) et que le fichier soit importé en tant que Table des nœuds (2) puis cliquer sur Suivant > (3). Cette procédure est à suivre dans le cas d'un fichier csv.



Enfin vérifier les informations comme dans l'exemple ci-dessous, puis cliquer sur Terminer.

L'étape Paramètres d'import est très importante : Gephi va reconnaître certaines des colonnes en fonction de leurs titres, mais vous aurez toujours à vérifier que le logiciel sera en mesure de comprendre la nature de vos données. Si vous travaillez avec des coordonnées, assurez-vous d'informer Gephi que les latitudes et les longitudes sont importées comme variable Double (et non pas en Integer).

Note importante : vous pouvez ne pas avoir de lien (comme c'est le cas avec les fichiers produits par le package Stylo). Dans ce cas là Gephi va reconnaître les nœuds à partir des arêtes : passez à l'étape suivante.

Edges

Pour les arêtes (edges), suivez la même procédure, mais avec le document des edges.

Importer feuille de calcul

Options générales

Choisissez un fichier CSV à importer :

tutoriels/01. Tuto Gephi/02. Exemple/Edges1.csv

Séparateur : Poin... (1) "Point-virgule" (2) (3)

En tant que table : Table des noeuds Table des liens

Encodage : UTF-8

Prévisualisation :

Source	Target	Type	Weight
376088951	17647430	Directed	1
376088951	32416061	Directed	1
376088951	550180187	Directed	2
376088951	28110685	Directed	3
376088951	870137221	Directed	3
550180187	376088951	Directed	2
550180187	17870064	Directed	2
550180187	320140078	Directed	1

Aide < Précédent Suivant > Terminer Annuler

Importer feuille de calcul

Paramètres d'import

New columns are created with the specified type.
A generated id is assigned if missing or already existing.
Edges need 'Source' and 'Target' columns with the id of the s
If no 'Type' column is provided, all edges will be directed.
If an edge already exists, attributes will be ignored, but their

Colonnes importées :

☒ Source String

☒ Target String

☒ Type String

☒ Weight Double

☐ Créer les noeuds manquants (1)

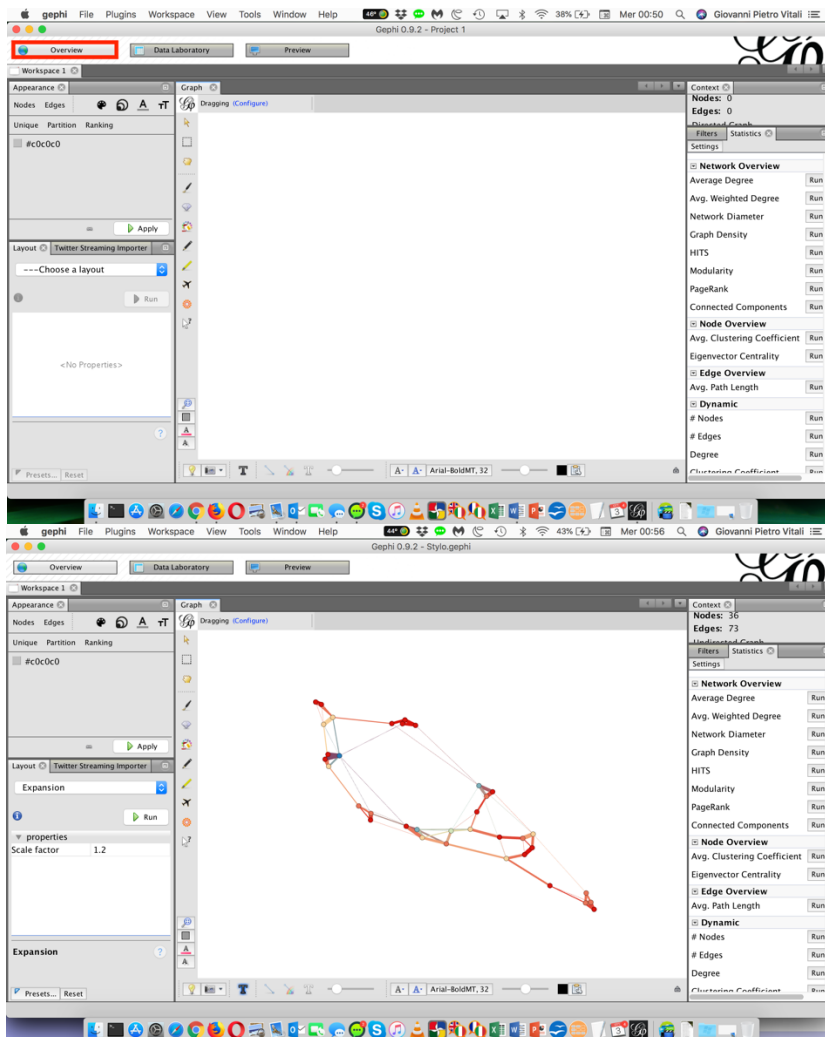
Aide < Précédent Suivant > Terminer Annuler

Dans les options générales d'importation, choisir que la séparation entre les colonnes en fonction du format de votre fichier² (1) et que le fichier soit importé en tant que Table des liens (2) puis cliquer sur Suivant > (3).

Puis vérifier les informations comme dans l'exemple ci-dessous, décocher la case Créer les nœuds manquant (1) – nous les avons déjà importés – puis cliquer sur Terminer.

2 N'hésitez pas à ouvrir le document pour contrôler si la séparation se fait avec des virgules, des points-virgules ou des tabulations.

Onglet 2 : Overview (vue d'ensemble)

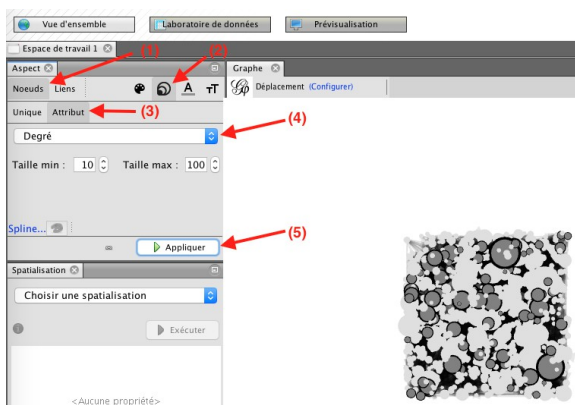


La section Overview permet de travailler les données à travers des filtres statistiques.

Les *nodes* et les *edges* peuvent être modifiés et coloriés pour en faciliter l'analyse par le chercheur.

Nous allons maintenant travailler dans l'onglet « Vue d'ensemble ». Le programme va produire une vue d'ensemble du graphique, avec une spatialisation aléatoire illisible.

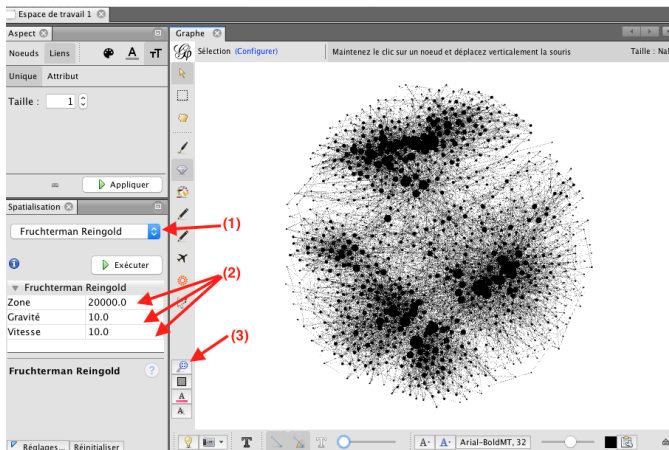
Taille des nœuds



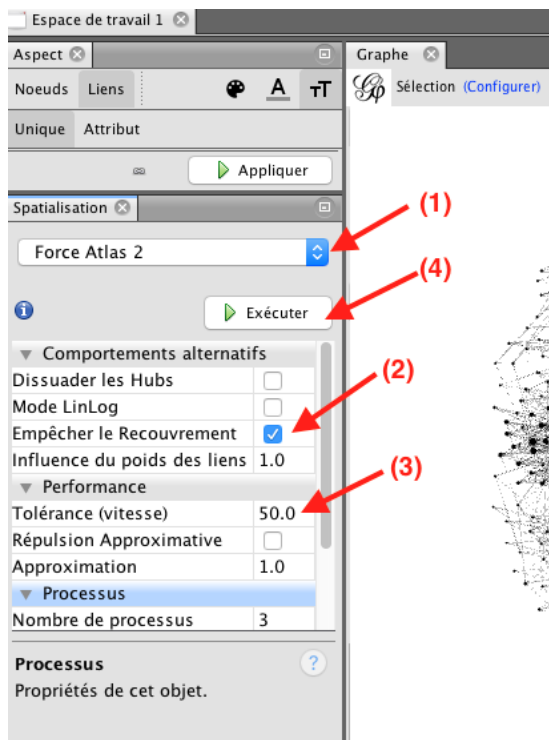
Pour donner aux nœuds une taille proportionnelle à leurs degrés (nombre de connexions). Dans le panneau Aspect de la colonne de gauche (en haut), sélectionnez Nœuds (1) et les cercles enchâssés (2), puis sélectionnez Attribut (3). Dans le menu déroulant (4), sélectionnez Degré et entrez la valeur minimale et maximale (nous proposons 10-100). Enfin cliquez sur Appliquer (5).

Spatialisation

Nous allons commencer par une spatialisation qui donne plus d'espace au graphique mais qui le maintient dans une aire définie. Pour cela, nous allons utiliser la visualisation de Fruchterman Reingold (1), avec les mêmes valeurs que dans l'exemple 20.000 - 10 - 10 (2). Cliquez à la fin sur exécuter. Cette visualisation dispose les nœuds de manière gravitationnelle (attraction-répulsion, comme si c'était des aimants).



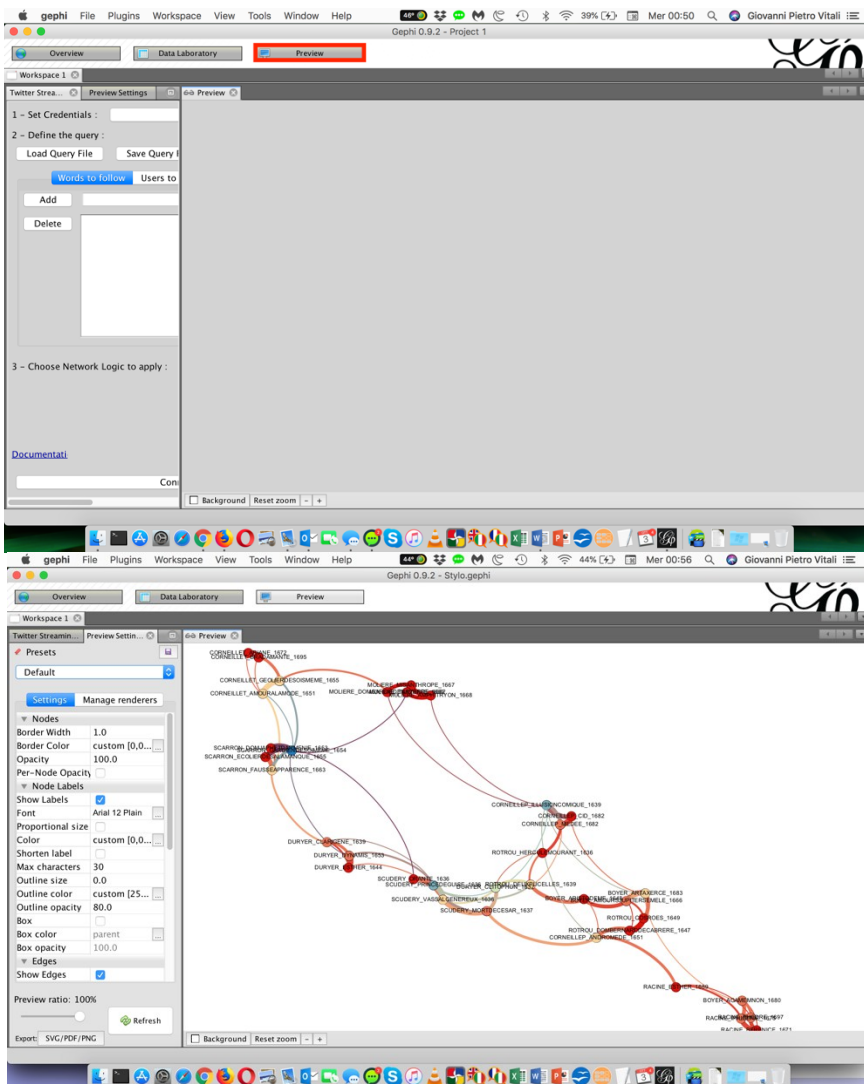
Utilisez la petite loupe bleue (en bas à gauche du panneau graphique) pour réajuster le zoom (3).



Ensuite, nous proposons d'utiliser la spatialisation « Force Atlas 2 » (1) – un autre algorithme de mise en page, pour disperser les groupes et laisser de l'espace autour des nœuds les plus importants. Attention, les paramètres que vous entrez modifient considérablement l'apparence finale. Nous proposons de cocher Empêcher le Recouvrement (1) et modifiez Tolérance (vitesse) en la mettant à 50.0. Laissez la fonction se développer en cliquant sur Exécuter (4), jusqu'à ce que le graphique soit stabilisé, puis cliquez sur Arrêter.

Nous pouvons appliquer Force Atlas 2 directement, sans appliquer Fruchterman Reingold avant, mais comme la « disposition aléatoire » initiale est une mise en page... aléatoire, il est préférable de modifier le réseau avant de le transmettre à un puissant algorithme de force.

Onglet 3 : *Preview* (prévisualisation)

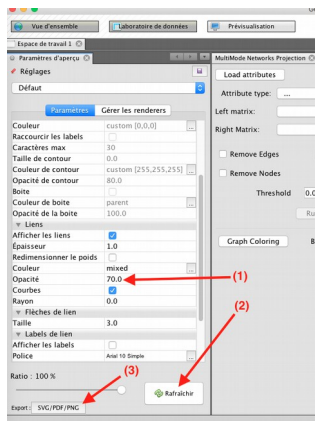


La section *preview* permet de travailler les données graphiquement et de les préparer pour l'exportation.

Dans cette section il est possible de transformer le graph pour le rendre plus compréhensible aux utilisateurs.

Finaliser le graphique

Passez dans le menu *Prévisualisation* pour arranger finaliser le résultat avant l'exportation. Contrairement aux étapes précédentes, la modification des paramètres dans ce menu est réversible et n'affecte pas la structure du graphique.



Au bas de cette colonne d'aperçu, vous trouvez un lien d'exportation (3). Notez que l'exportation au format .png produit une figure avec une résolution médiocre. Vous voudrez peut-être opter pour .svg ou .pdf, qui ont l'avantage d'être modifiables par votre propre logiciel de traitement d'image (je recommande le programme *open source* inkscape pour manipuler des fichiers .svg).