Data Collection and Visualization

An introduction to TAGS and Gephi

Tuja Khaund, COSMOS

Outline

1. Data Collection using TAGS
    a. Background
    b. How it collect data?
2. Data Visualization using Gephi
    a. Background
    b. Data
    c. How to visualize data?

TAGS - an Introduction

TAGS - Background

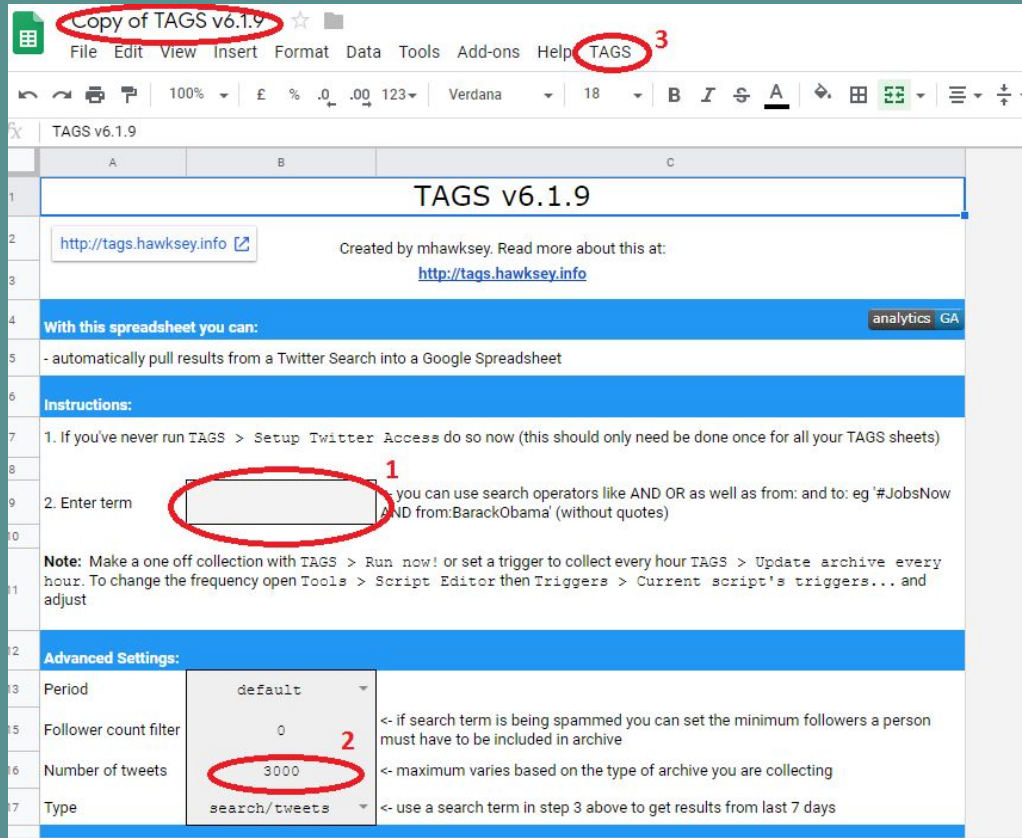
- TAGS stands for **T**witter **A**rchiving **G**oogle **S**heets
- Developed as a ‘hobby project’ by Martin Hawksey in june 2010
- Six major versions have been released since its origin. Learn more about its history here.

TAGS - How to collect data?

- Requirements:
  - A Twitter account, and a Google account for authentication
- Getting started:
  1. Visit the url https://tags.hawksey.info/ to start
  2. Click on '**Get TAGS**' to proceed; then click on '**TAGS v6.1**'
  3. Click on '**Make a copy**' to get started
- The entire data collection process is also available here.

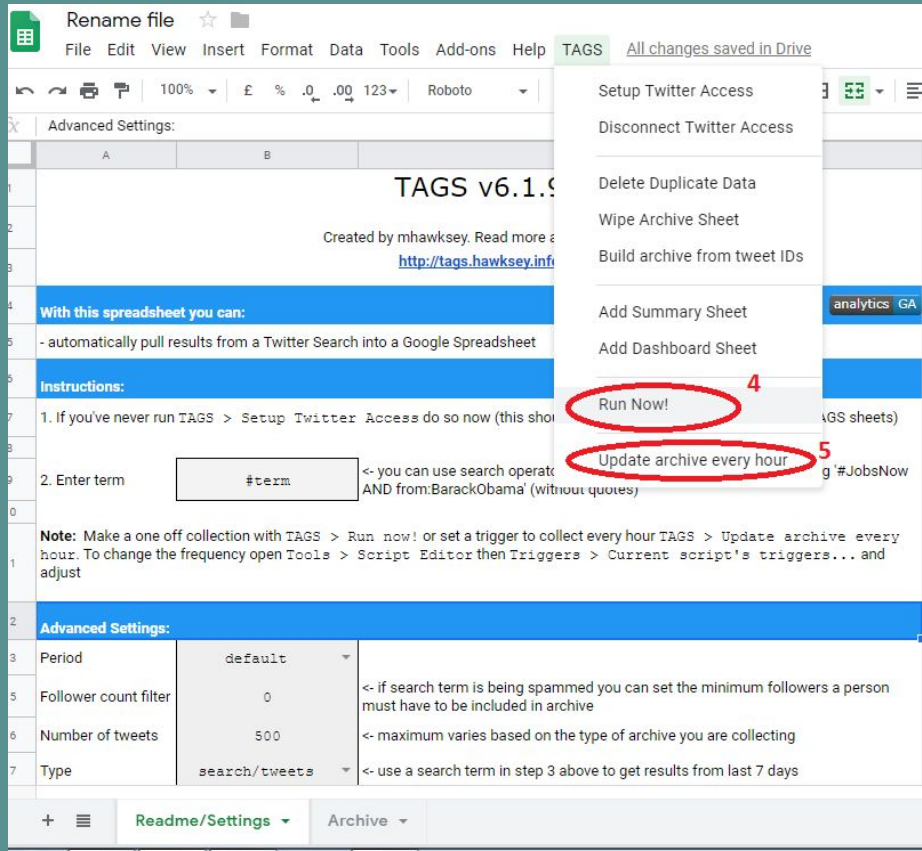
TAGS - How to collect data?

- Steps 1, 2 and 3:



TAGS - How to collect data?

- Step 4, 5:
  - The output can be found in the '**Archive**' sheet from bottom left



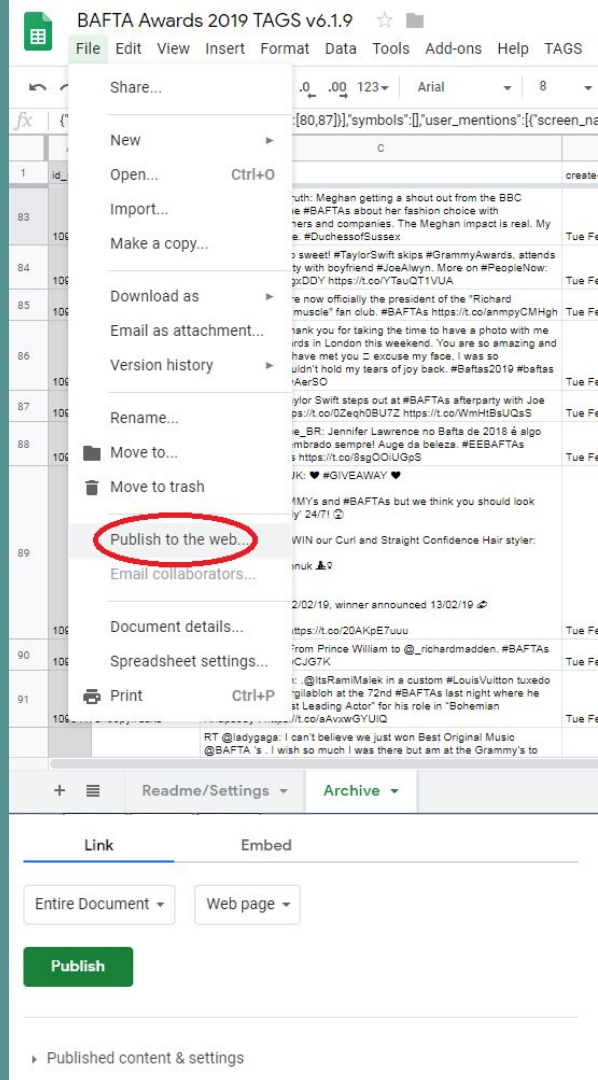
TAGS - How to collect data?

- Data Snippet:
  - The dataset is available to view here.

id_str	from_user	text	created_at	time	geo_coordinates	user_lang	in_reply_to_user_id_	in_reply_to_screen_r	from_user_id_str	in_reply_to_status_id	source	profile_image_url
109512	toogoodyy	RT @royallymegtruth: Meghan getting a shout out from the BBC commentator at the #BAFTAs about her fashion choice with sustainable designers and companies. The Meghan impact is real. My girl ain't even there. #DuchessofSussex	Tue Feb 12 00:43:48	12/02/2019 00:43:48		en			1594557132		<a href="http://twitter.	http://pbs.twimg.com/
109512	perez_lorenzo1	RT @peopletv: So sweet! #TaylorSwift skips #GrammyAwards, attends #BAFTAs afterparty with boyfriend #JoeAlwyn. More on #PeopleNow: https://t.co/82FsugxDDY https://t.co/YTauQT1VUA	Tue Feb 12 00:43:31	12/02/2019 00:43:31		es			941414241449283584		<a href="http://twitter.	http://pbs.twimg.com/
109512	MariaBelenn_3	RT @enews: We're now officially the president of the "Richard Madden's left jaw muscle" fan club. #BAFTAs https://t.co/anmpyCMHgh	Tue Feb 12 00:43:17	12/02/2019 00:43:17		en			1029870067083866115		<a href="http://twitter.	http://pbs.twimg.com/
109512	louisaharriet	@MRodOfficial Thank you for taking the time to have a photo with me at the BAFTA awards in London this weekend. You are so amazing and I am honoured to have met you ☐ excuse my face, I was so overwhelmed. Couldn't hold my tears of joy back. #Baftas2019 #baftas https://t.co/wqHSyAerSO	Tue Feb 12 00:43:16	12/02/2019 00:43:16		en	51219631	MRodOfficial	50399071		<a href="http://twitter.	http://pbs.twimg.com/
109512	lizandersoo	RT @billboard: Taylor Swift steps out at #BAFTAs afterparty with Joe Alwyn: see pic https://t.co/0Zeqh0BU7Z https://t.co/WmHtBsUQsS	Tue Feb 12 00:40:43	12/02/2019 00:40:43		es			1559182224		<a href="http://twitter.	http://pbs.twimg.com/
109512	selenativist	RT @JenLawrence_BR: Jennifer Lawrence no Bafta de 2018 é algo que merece ser lembrado sempre! Auge da beleza. #EEBAFTAs #BAFTA #BAFTAs https://t.co/8sgOOiUGpS	Tue Feb 12 00:40:42	12/02/2019 00:40:42		pt			3289935290		<a href="http://twitter.	http://pbs.twimg.com/

TAGS - Visualization

- Walkthrough:
  - Important! - Make sure the file sharing is set to **public**
  - You will need to publish your work on the web
  - Click on '**TAG Explorer**' to visualize your data
  - You can view a sample visualization here.



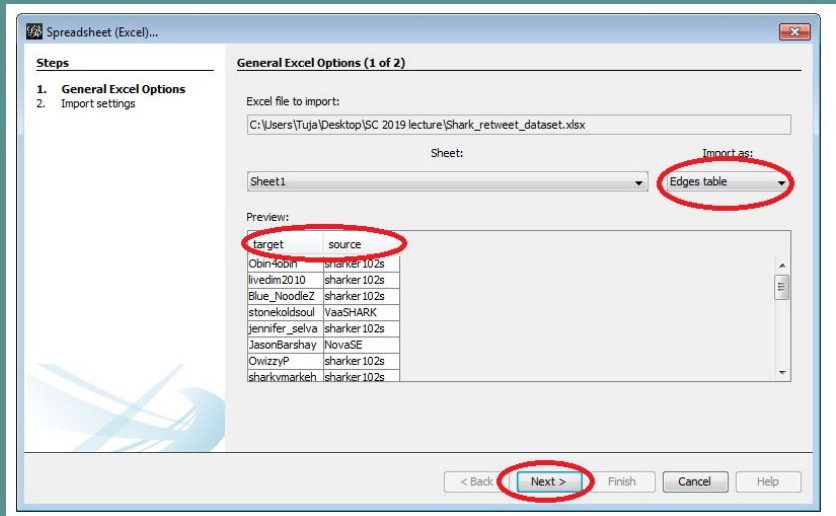
TAGS - Visualization

- Pros:
  - Visualization can be done with just one click
  - Can get better insights without data preprocessing
- Cons:
  - Limited to only three types of networks, reply, mention and retweets
  - Cannot make any visual changes
  - Hard to visualize big datasets

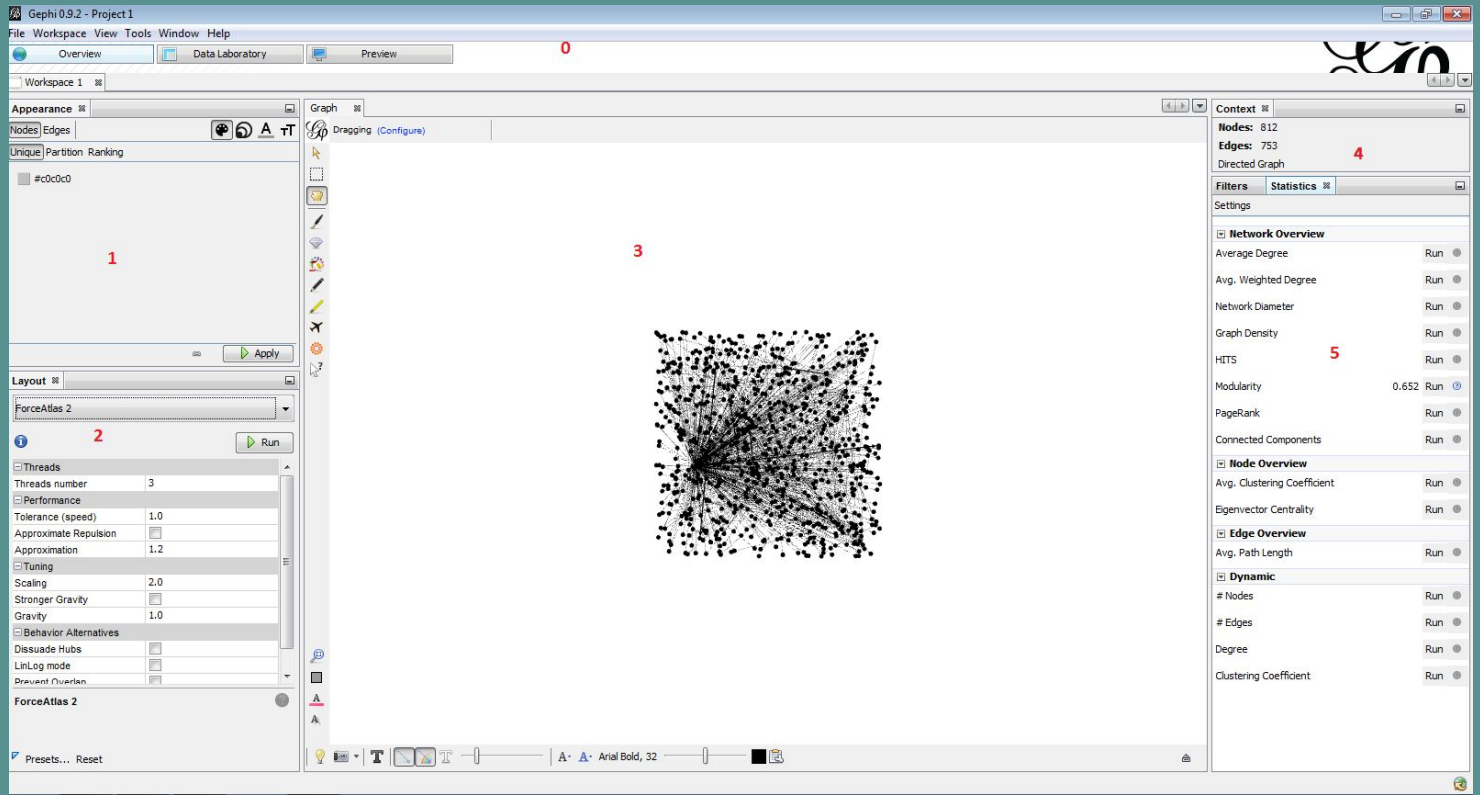
Gephi - a graph visualization tool

Gephi - Install and Import

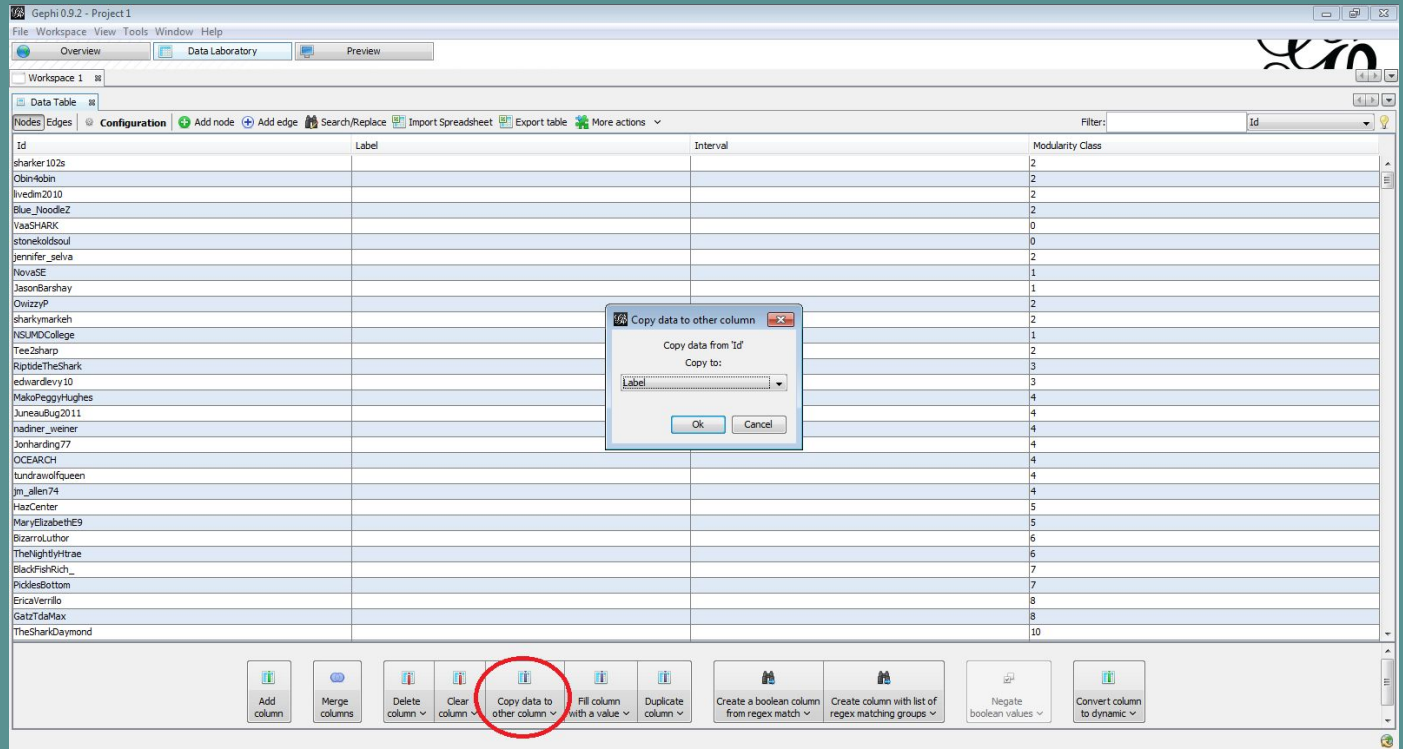
- Step 1: Get set up
  - Download Gephi, install it, open it up and start a new project.
- Step 2: Import a Spreadsheet
  - Input file needs two columns, “Source” and “Target.”
  - Click *File > Import spreadsheet*
  - Select ‘directed’ from the next screen



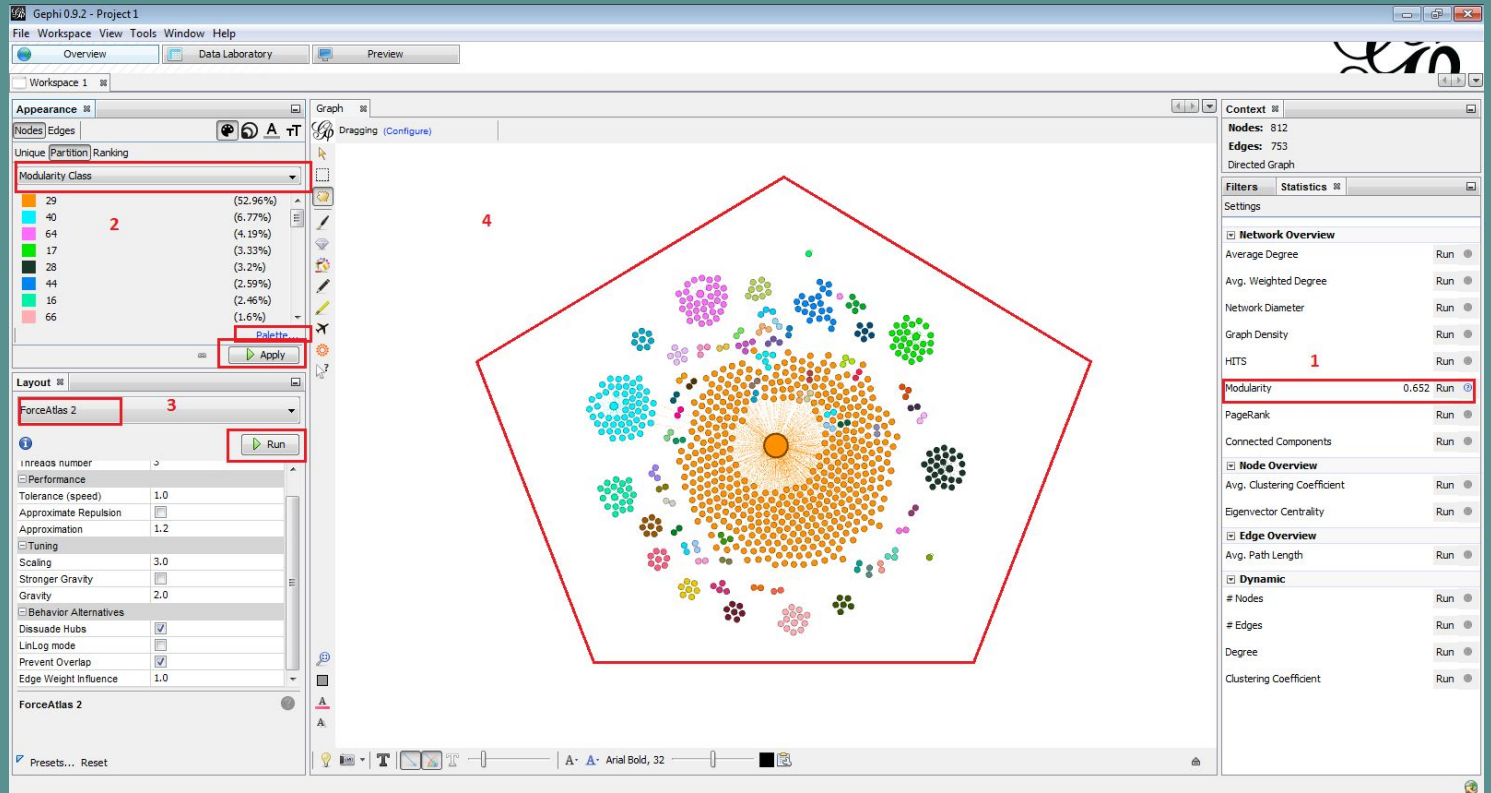
Gephi - understand the User Interface



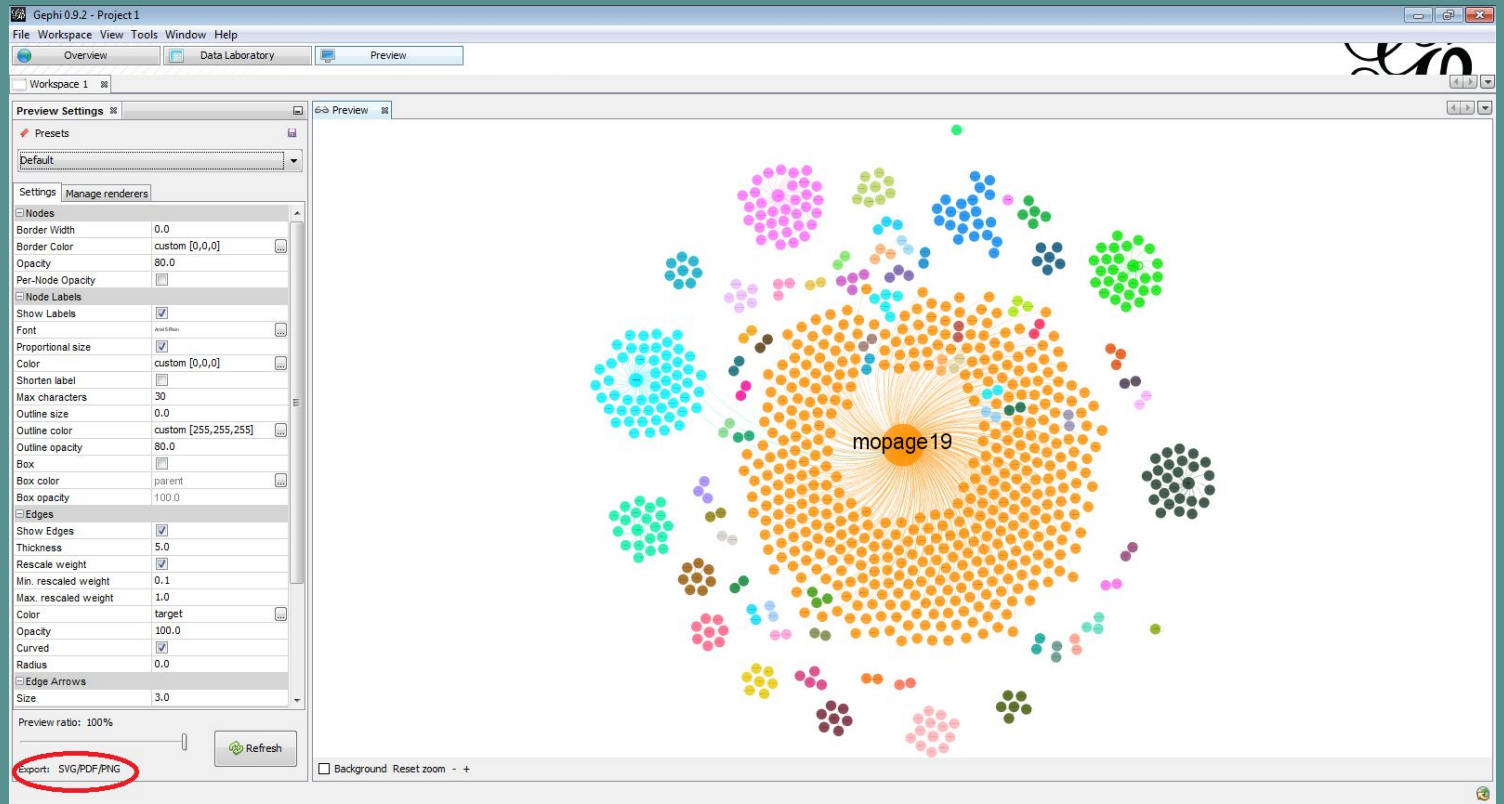
Gephi - creating labels



Gephi - How to visualize?

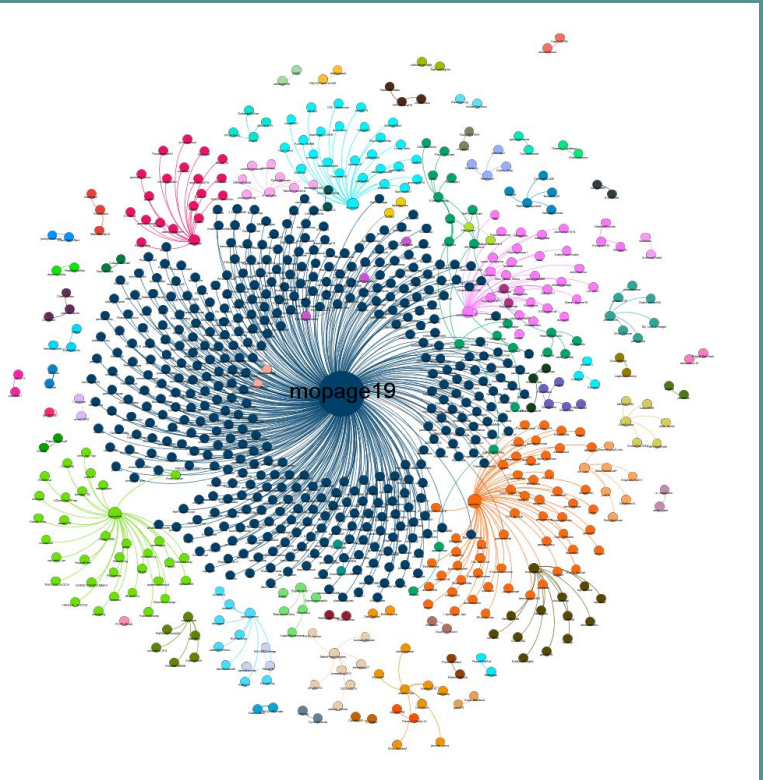


Gephi - How to visualize?



Sample Data and Visualization

	A	B	C	D
1	**target**	**source**		
2	Obin4obin	sharker102s		
3	livedim2010	sharker102s		
4	Blue_NoodleZ	sharker102s		
5	stonekoldsoul	VaaSHARK		
6	jennifer_selva	sharker102s		
7	JasonBarshay	NovaSE		
8	OwizzyP	sharker102s		
9	sharkymarkeh	sharker102s		
10	NSUMDCollege	NovaSE		
11	Tee2sharp	sharker102s		
12	edwardlevy10	RiptideTheShark		
13	JuneauBug2011	MakoPeggyHughes		
14	nadiner_weiner	MakoPeggyHughes		
15	Jonharding77	MakoPeggyHughes		
16	OCEARCH	MakoPeggyHughes		
17	tundrawolfqueen	MakoPeggyHughes		
18	jm_allen74	MakoPeggyHughes		
19	MaryElizabethE9	HazCenter		
20	TheNightlyHtrae	BizarroLuthor		



Thank You!

Questions?