Esercitazione su Gephi

Rachele Sprugnoli – <u>rachele.sprugnoli@unicatt.it</u>

Centro Interdisciplinare di Ricerche per la Computerizzazione dei Segni dell'Espressione (CIRCSE)



GEPHI

- Gephi è "a visualization and exploration software for all kinds of graphs and networks"
 - visualizzazione e analisi delle reti
 - calcolo integrato di metriche complesse
 - stand-alone
 - open source, gratuito
 - https://gephi.org/



GEPHI: FORMATO DATI

 2 file CSV (si può creare partendo da un file Excel o Numbers): uno per i nodi e uno per gli archi

NODI - NODES



- La prima riga con le etichette delle colonne è obbligatoria
- L'unica colonna obbligatoria è quella dell'identificativo
- La colonna con l'identificativo deve avere "Id" come etichetta
- La colonna coi nomi dei nodi deve avere come etichetta "Label"
- Gli attributi si aggiungono con altre colonne

GEPHI: FORMATO DATI

 2 file CSV (si può creare partendo da un file Excel o Numbers): uno per i nodi e uno per gli archi

ARCHI - EDGES

Source	Target	Туре
4	1	Directed
5	2	Directed
6	2	Directed
7	2	Directed
8	2	Directed

- La riga con le etichette è obbligatoria
- Due colonne sono obbligatorie: Source e Target
- La colonna Type può avere solo due valori: Directed o Undirected
- Importando solo la tabella degli archi, si creano automaticamente i nodi corrispondenti

ANALISI DELLE INTERAZIONI IN UN'OPERA TEATRALE NODI=PERSONAGGI

- Aprite Gephi (doppio click sull'icona)
- 2. Cliccate su New Project
- 3. Cliccate su Data Laboratory
- 4. Cliccate su Import Spreadsheet
- Scegliete il file manzoni-adelchi.csv (da https://dracor.org/ita): selezionate separatore "comma" e importate come "Edge Table"
- 6. Cliccate su Successivo e poi su Fine. Gephi potrebbe chiedere se volete aggiungere la tabella a un nuovo progetto o aggiungerla alla tabella dati esistente: scegliete questa seconda opzione
- 7. Nella tabella Nodes, in basso, cliccate su "Copy data to other column", scegliete "Id" e copiate su "Label" poi su "ok": così abbiamo le etichette dei nodi

- 8. Nella tabella Nodes, in basso, cliccate su "Add Column" e aggiungete come titolo "Gender": aggiungete il genere dei personaggi (M/F)
- 9. Cliccate su Overview
- 10. Nella sezione Layout, scegliete "ForceAtlas2" e cliccate su "Run"
- 11. Modificare le opzioni "Scaling = 200", "Dissuade Hubs" e "Prevent Overlap" e cliccate su "Run"
- 12. Cliccate sulla "T" in basso sotto la rete per far apparire le etichette dei nodi
- 13. Cliccate su Preview per vedere la rete
- 14. Tornate su Overview: nella sezione Appearance (a sinistra) scegliete "Partition", sotto "Choose an Attribute" scegliete "Gender", poi cliccate su "Apply"

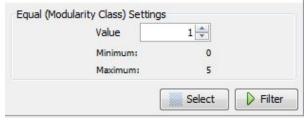
- 15. Nella sezione Appearance scegliete "Size" ♠ ♠ ♣ т poi "Ranking", sotto "Choose an Attribute" scegliete "Degree", poi cliccate su "Apply"
- 16. Nella sezione Statistics, alla destra della rete, cliccate su "Modularity" e su "Avg. Path Length"
- 17. Tornando a sinistra nella sezione Appearance sotto "Ranking" scegliete "Betweenness Centrality" poi "Apply": come cambiano i nodi?
- 18. Nella sezione Appearance sotto "Partition" scegliete "Modularity Class" e poi "Apply": che gruppi possiamo identificare?
- 19. Nella sezione Preview, cliccate su Refresh per aggiornare la visualizzazione della rete

ANALISI DEL DATASET DELLA TATE GALLERY NODI=ARTISTI E MATERIALI

- Aprite Gephi (doppio click sull'icona)
- 2. Cliccate su Open Graph File...
- 3. Selezionate e aprite il file Tate-1800.gephi
- 4. Cliccate su Modularity nella sezione Statistics e poi su ok: quante comunità vengono rilevate?
- 5. Nella sezione Appearance, selezionate Partition sotto Nodes, scegliete Modularity Class dal menu a tendina e cliccate su Apply: cosa ci dicono le percentuali?

ANALISI DEL DATASET DELLA TATE GALLERY NODI=ARTISTI E MATERIALI

- Cliccate su Filters (accanto a Statistics) poi sul + accanto ad Attributes poi sul + accanto a Equal
- Fate doppio click su Modularity Class: apparirà il filtro Equal (Modularity Class) nella sotto-sezione Queries
- Modificate il Value nei Settings per vedere meglio le varie comunità e poi cliccate su Filter



9. Per rimuovere il filtro, cliccate col tasto destro su Equal (Modularity Class) nella sotto-sezione Queries e poi su Remove

ANALISI DEL DATASET DELLA TATE GALLERY NODI=ARTISTI E MATERIALI

- Alcune note
- i dati di partenza provengono dai file già usati per Palladio: ho creato un file csv da artwork_data.ods selezionando solo le colonne artist e medium per creare gli archi
- l'immagine nella sezione Preview si aggiorna in base al filtro attivato
- è possibile scaricare i fogli di calcolo corrispondenti a nodi e archi dalla sezione Data Laboratory e cliccando su Export table
- i dati nella sezione Data Laboratory cambiano a seconda del filtro attivato: per ottenere tutti i dati, disabilitare i filtri
- cosa notate se calcolate la Modularity per il file Tate-2000.gephi?

ESPORTAZIONE

 Nella sezione Preview, cliccare su Export SVG/PDF/PNG: per ogni tipo di formato ci sono delle possibili opzioni da scegliere



 Nel menu File, scegliere Save as per salvare in formato Gephi o scegliere Export per altri formati

APPROFONDIMENTI

Tutorial Gephi:

https://gephi.org/users/

http://www.martingrandjean.ch/gephi-introduction/

http://www.kateto.net/wp-content/uploads/2012/12/COMM645%20-%20Gephi%20Handout.pdf

ALTRI STRUMENTI:

- The Vistorian: https://vistorian.net/
- Cytoscape: https://cytoscape.org/
- Social Network Visualizer: https://socnetv.org/



GRAZIE!

Email: rachele.sprugnoli@unicatt.it

Twitter: @RSprugnoli



