Project Task 2025

In this project, you should ask and answer an interesting question about musical artists based on a given data set fro secondhand song.org.

Suggestion: You might ask *Which musical artist is the most influential musician between 1960 and 2020?*

Use Network Science methods to contribute to answering the question.

Datasets:

The datasets are based on dump of the secondhandsong.org project. They describe for many songs, recorded later than 1960, the original artist, and artists that covered the song (before 2018)

The following files can be found on myStudy.

- Artists: relevant musicians / bands
- Originals: songs, that have been covered at least once
- Covers: cover-song with references to original song
- Releases: LPs, CDs etc.

Suggested Steps:

1. Choose a network models (see Brandes et al, 2013) that allows you to answer your question. Explain your network concept and justify your decisions.
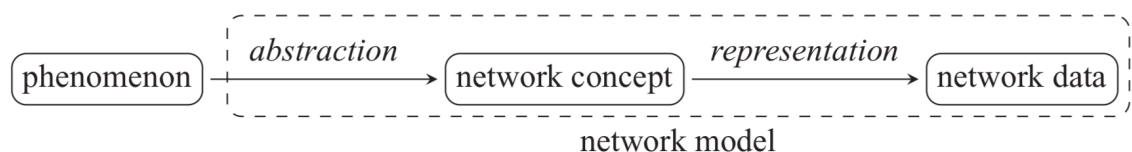

Fig. 1. The elements of network models.

2. Choose the network data: which data do you want to use? (This may - but does not have to - involve subsets of the given dataset)
3. Construct a network using Python and/or the Neo4J graph database.
4. Describe the network by the standard graph metrics (number of vertices / edges, number of components, average distance, ...)
5. Apply appropriate centrality measures or clustering methods that help you to answer the question. Justify your choice of metrics and their significance in the respective network model.

6. Create suitable visualisations of graph excerpts that you can use to explain the construction and results of your analysis.
7. Answer your question!
8. Name any limitations of your investigation.

Deliverables:

Upload the following documents until 31.8.2025 via myStudy:

- Project documentation of 10-12 pages per group (for each team member, indicate which text sections they contributed to.)
- In case of additional datasets -> a link to the data)
- All Python sources in the form of one (or more) executable notebooks
- All Cypher scripts from your Neo4J databases