# **Network Analysis Project Title**

Andrea D’Arpa, Degree Programme, ID number (matricola)

Chloe Papadopoulos, Digital Humanities and Digital Knowledge, ID number (matricola)

Maddalena Ghiotto, Digital Humanities and Digital Knowledge, 1030096

### **Context**

The general field in which our research can be situated is Film Studies, a sector concerned with exploring the narrative, artistic, cultural, economic, and political implications of the cinema.

In particular, we are interested in adopting a postcolonial perspective as a critical approach to the analysis of production, theoretical frameworks, contexts and creation of films.

To do so, our study will consider the history of films and the reciprocal influence between USA and East Asian cinema production, hence adopting the tools and methods of citation analysis studies.

### **Problem and Motivation**

The problem we aspire to address within this study is the power asymmetries underlying the ever-present vision of history of Cinema as Hollywood-centered, as well as how eastern cinema has been incorporated in western cinema, often without fully acknowledging the sources, hence undermining the role of the foreign artistic influence; or how it affected the representation of the Other, shaping the perception and history of foreign identities.

As a compass to better navigate this field of studies we draw knowledge from two main texts:

* *Postcolonial Cinema Studies*, edited by Sandra Ponzanesi, and Marguerite Waller, Taylor & Francis Group, 2011. *ProQuest Ebook Central*
* *East Asian Cinemas, Exploring Transnational Connections on Film,* edited by Leon Hunt & Leung Wing-Fai, I.B Tauris & Co Ltd, New York, 2008

Moreover, we are taking inspiration from Spitz and Horvàt’s Network Analysis study *Unraveling Cinematic Citation[[1]](#footnote-1)* where light is shed « on the potential of a systematic quantitative investigation based on cinematic influences in identifying the most inspiring creations in world cinema».[[2]](#footnote-2) Their main contribution was to create for the first time a movie rank based on the historical importance of movies, beyond traditional measures of success, such as box-office revenue and critical acclaim, which are external to the craft of the movie itself.

Nonetheless, their result is not critical to underrepresentation of East Asian Cinema and their results, as one could expect, reflect a western narration of cinema history.

Our specific contribution will be in attempting to:

* Produce a long-term impact analysis that specifically ranks East Asian films according to their influence on American cinema.
* Produce a general long-term impact film rank where the representation of East Asian cinema is weighted and balanced against the overrepresentation of American Cinema.
* Individuate central figures in transnational film production collaboration in order to assess how production with transnational teams relate to the success of the film, and to what extent it agrees with the most famous ranks of movie’s success

### **Datasets**

The project’s data is acquired by the Internet Movie Database, which provides datasets in tab-separated values format, freely accessible on their dedicated [website](https://datasets.imdbws.com/) for personal and non-commercial use.

We used the subset title.akas.tsv.gz[[3]](#footnote-3), which contains the following information for titles:

* titleId (string) - a tconst, an alphanumeric unique identifier of the title
* ordering (integer) – a number to uniquely identify rows for a given titleId
* title (string) – the localized title
* region (string) - the region for this version of the title
* language (string) - the language of the title
* types (array) - Enumerated set of attributes for this alternative title. One or more of the following: "alternative", "dvd", "festival", "tv", "video", "working", "original", "imdbDisplay". New values may be added in the future without warning
* attributes (array) - Additional terms to describe this alternative title, not enumerated
* isOriginalTitle (boolean) – 0: not original title; 1: original title

Regarding information on movie citations, it appears that the dataset movie-links.list used in our reference paper had been discontinued by IMDb, as we could not find an updated version on their platform. Luckily, we were able to find it in another [repository](http://givingsense.eu/opendata/data/imdb-movies/database/) in a text file format.

**What tools did you use 1) to handle (store, manipulate) the data and 2) to compute measures on the data?**

### **Measures**

What measures did you apply (brief explanation of how they work)? How do they relate to the intent of the study? Why are they relevant?

1. Andreas Spitz, Emőke-Ágnes Horvát, *Measuring Long-Term Impact Based on Network Centrality: Unraveling Cinematic Citations,* published in Plos One, 8 October 2014 (<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0108857#references> ). [↑](#footnote-ref-1)
2. *Ibidem.* [↑](#footnote-ref-2)
3. 1. The Internet Movie Database (IMDb). Alternative interfaces. <http://imdb.com/interfaces/> [↑](#footnote-ref-3)