Gephi ist eine Visualisierungssoftware für alle Arten von Grafiken. Es ist Open Source und kostenlos. In unserem Projekt haben wir Gephi für einen Knowledge Graphen benutzt. Für die Software haben wir uns entschieden, da sie sehr einfach zu bedienen ist. Man kann hier sehr gut zwei oder mehrere Werte miteinander verbinden welches für die Erstellung eines Knowledge Graphen wichtig ist und für uns auschlaggebend war. Für Gephi liest man in den meisten Fällen eine CSV Datei ein, die die Werte beinhaltet, die man miteinander verbinden will. Daraufhin erstellt die Software die Verbindungen, wodurch sich ein Knowledge Graph ergibt.

In unserem Gephi Projekt gibt es Knoten und Kanten. Bei den Knoten handelt es sich um 500 Filme aus mehreren Genres. Die verschiedenen Genres sind folgende:

Action, Abenteuer, Drama, Biografie, Animation, Comedy, Krimi und Horror.

Zudem befinden sich in unserer Input Datei, in der sich die 500 Filme befinden, ebenfalls die dazugehörigen Genres und die Regisseure der einzelnen Filme. Die Input Datei haben wir verwendet, um mithilfe unseres Automatisierungscodes, welches wir mit Python geschrieben haben, unsere „Unsere\_Datei\_movieID\_serial\_ref.csv“ erstellen zu können. In der „Unsere\_Datei\_movieID\_serial\_ref.csv“ Datei befinden sich zwei Spalten. Die Spalten Source und Target sagen aus, welche Filme in unserem Knowledge Graphen miteinander verknüpft werden sollen. Es werden alle Filme in dieser Datei miteinander verknüpft, welche dasselbe Genre und denselben Regisseur aufweisen. Diese Verknüpfungen wurden mithilfe des Automatisierungscodes erstellt. Die Python Codes befinden sich in den zwei Python-Dateien „movie\_references.py“ und „reorg\_df\_index\_serial.py“. Nach der Erstellung der zwei CSV-Dateien wurden diese in Gephi über das Datenlabor importiert. Der Knowledge Graph zeigt uns nun die Filme an, die denselben Regisseur und dasselbe Genre aufweisen. Bei Kanten mit einem dicken Pfeil handelt es sich um die Verbindung von Filmen mit demselben Regisseur. Die dünneren Linien bzw. Verbindungen stehen für das Genre. Es handelt sich bei unserem Graphen um einen gerichteten Graphen mit 47558 Kanten. Das bedeutet wir haben 47558 Verbindungen bei 499 Knoten. Beim Layout haben wir uns für das „Fruchterman Reingold“ Layout entschieden, da dieses Layout ein schönes Cluster vorweist.