# Knowledge Graph Analysis

## Exercise Sheet 2

### Dr. Hamed Shariat Yazdi, Prof. Jens Lehmann

November 8, 2018

## 1 In Class

1. Answer these questions on graph databases and the labeled property graph model!
   - ▷ What is the difference between *Labels* and *Properties* in labeled property graphs?
   - ▷ What are disadvantages and advantages of relational databases compared to graph databases?
   - ▷ Can we use an undirected relationship in creating a database in Neo4j? And what about querying in Cypher?
   - ▷ What could be nodes, labels, properties & relationships in the statements below:
     *"Mark is 24 years old. He is a programmer. He has a cat named "Fifi". His wife is Maria, a dentist."*

2. Look at these two modeling strategies for movies people like, and say which one you would choose and why?
   - ▷ In model 1, there are 10 persons and 10 movies, each node is labeled by its name (in case of being a person) or title (in case of being a movie).
   - ▷ In model 2, there are again 10 persons and 10 movies, each node is labeled by its concept (i.e. 10 nodes are labeled Person and 10 are labeled Movie) and each node has its name/title as a property.

3. Create a property graph covering the statements below.
   - ▷ Anne likes tennis and has a 3-years-old boy named Nick.

    ▷ Friedhelm likes literature and Jacob plays tennis.

    ▷ Helen likes Friedhelm, Jacob likes Helen and Friedhelm likes Helen.

    ▷ Helen has known Anne since 2001, Anne has known Alex since 1998, Friedhelm has known Jacob since 2000.

4. Query the answer to all questions below based on the depicted labeled property graph. In brackets, we list what you should return.
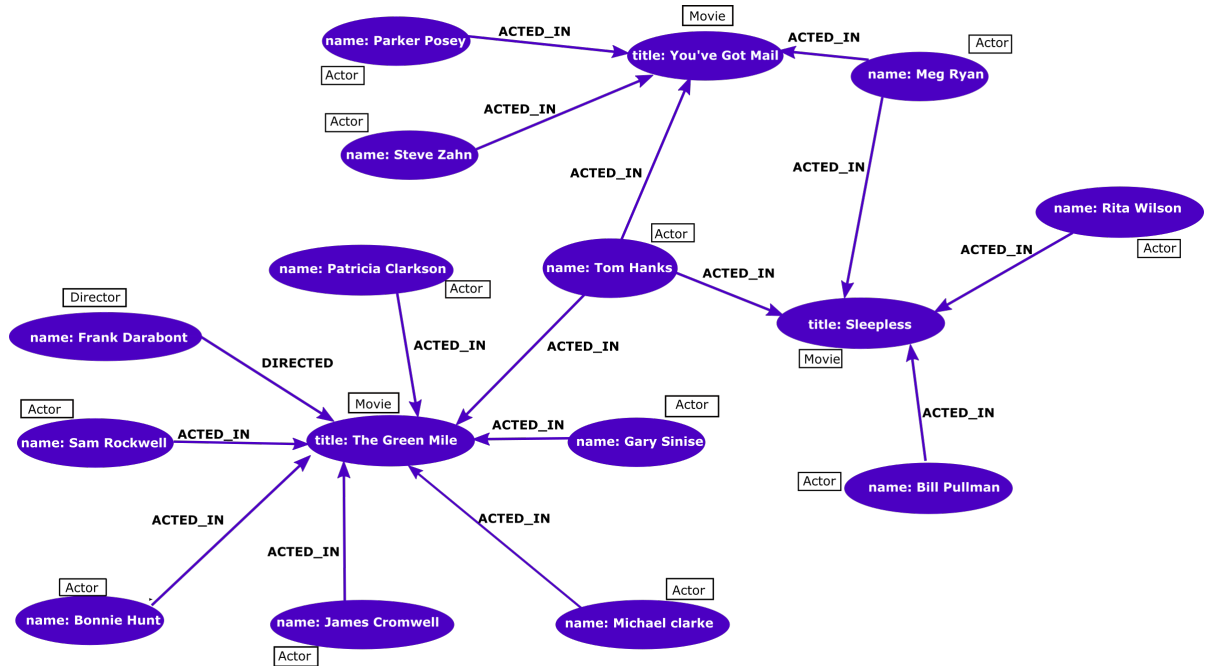


    ▷ Who is the director of "*The Green Mile*"? (name)

    ▷ List all the actors of "*The Green Mile*" and sort them in ascending order! (names)

    ▷ How many actors acted in "*The Green Mile*"? (number)

    ▷ List all actors that have co-acted with "Tom Hanks" along with the corresponding movie titles! (list of name + title)

    ▷ How many persons have been co-actors with Tom Hanks? (number)

5. Transform the RDF model below (based on the last exercise) to a property graph model!

```
@prefix ex: <http://example.org/movies/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

ex:m1    rdf:type ex:Movie;
```

```
        ex:genre ex:Drama;
        ex:year "2006"^^xsd:gYear;
        rdfs:label "Marie Antoinette";
        ex:country ex:USA;
        ex:director ex:p1;
        ex:actor ex:p2 .
ex:p1   rdf:type ex:Director;
        foaf:familyName "Coppola";
        foaf:givenName  "Sofia";
        ex:birthYear "1971"^^xsd:gYear .
ex:p2   rdf:type ex:Actor;
        foaf:familyName "Dunst";
        foaf:givenName  "Kirsten";
        ex:birthYear "1982"^^xsd:gYear .
```

## 2 AT HOME

Please install Neo4j (https://neo4j.com) for the exercises.

1. Take your model from Q3. Use Neo4j and command "CREATE" in Cypher create a graph. Illustrate the graph Using "MATCH (n) RETURN n" and see how it looks.

2. Query the answer of all questions below using the model from Q3:
   ▷ Anne wants to find someone who can play tennis with her.
   ▷ Find people who knew each other already in 1990s.
   ▷ Helen can't stand children, so with whom may she lose her contact?
   ▷ Who may get into a relationship based on mutual interest?

3. Using Neo4j create the graph of movies "Apollo 13" and "Cast Away" in the movie database provided by Neo4j and list all casts of these movies. You should use the "*:play movie-graph*" command to reach the database. Execute all Cypher queries of Q4.