**Large Scale and Multi-Structured Data Bases**
**University of Pisa**
**Academic Year 2019-2010**
**Workgroup Task 3**

**Title: Designing and Implementing an application which interacts with a Graph-based DB**

**Task 3.1: Dataset identification and design of the application and of the DB**

- The workgroup ***must first design*** a complete application which manages data that can be stored in a graph-based DB.
- To this aim, a suitable dataset along with appropriate requirements must be identified.
- Then, the UML diagrams of the use cases and of the analysis classes diagram must be drawn.
- Finally, a snapshot of the graph, in terms of entities, relationships and their main properties must be provided.

**Constrains:**

- CRUD operation must be implemented
- A couple of typical "on-graph" queries, must be defined (both domain-specific queries and graph-centric queries.).
- Some administration use cases must be defined

**Suggestions**:

- Use the attached examples for taking inspiration for the queries. It is forbidden to use the Movie dataset.
- The task can be embedded (or using the same dataset) into Task1 or Task2. To this aim, the requirements, the use cases diagram and, if necessary, the analysis class diagram must be appropriately extended.

**Task 3.2: Implementation and test of the application**

Once the Design document will be approved by the instructor, the workgroup will be allowed to start with the implementation.

A final document must be produced including:

- A description of the main modules of the software
- A description of the performed tests
- A short manual of usage of the application

At the end of the task, all the artifacts (reports, code, database dump and executable files) must be uploaded *only by the reference person* of the group. Avoid multiple uploads.
Groups *may use git archive*. In this case, the reference person must specify the address to download the repository. Please upload documentation always on the e-learning platform.

***Deadline: January 1, 2020.***