

Big Data Mangement - A.A. 2013/14
Written Exam of **June 26, 2014**

Exercise 1

Describe the main characteristics of the Star Schema and of its variant called Snowflake Schema. Examples are welcome.

Exercise 2

(a) Write (in any format you like) an RDF/RDFS model representing the following statements about URIs Player, Team, PlaysIn, IsCaptainOf, Totti, Destro, Roma.

1. Player and Team are classes;
2. PlaysIn and IsCaptainOf are properties;
3. the domain of both PlaysIn and IsCaptainOf is Player;
4. the range of both PlaysIn and IsCaptainOf is Team;
5. IsCaptainOf is a sub-property of PlaysIn;
6. Totti and Destro are Players;
7. Roma is a Team;
8. Both Totti and Destro play in Roma;
9. Totti IsCaptainOf Roma.

(b) Write SPARQL queries corresponding to the following requests: (b1) “return all pairs of players playing in the same team”; (b2) “return all tuples composed by a player, the team he plays in, and optionally the team he is captian of”.

Exercise 3

NoSQL databases, among various properties they have, are defined as *schemaless*. Explain the meaning of this term, describe how “schemalessness” is achieved in the various NoSQL data models studied in this course, and discuss advantages and drawbacks that come together with this property.

Exercise 4

Describe the map-reduce computing framework making it clear which is the role of the main functions it consists of. Examples are welcome but pseudo-code is not needed (it is sufficient to show input and outputs of functions).