

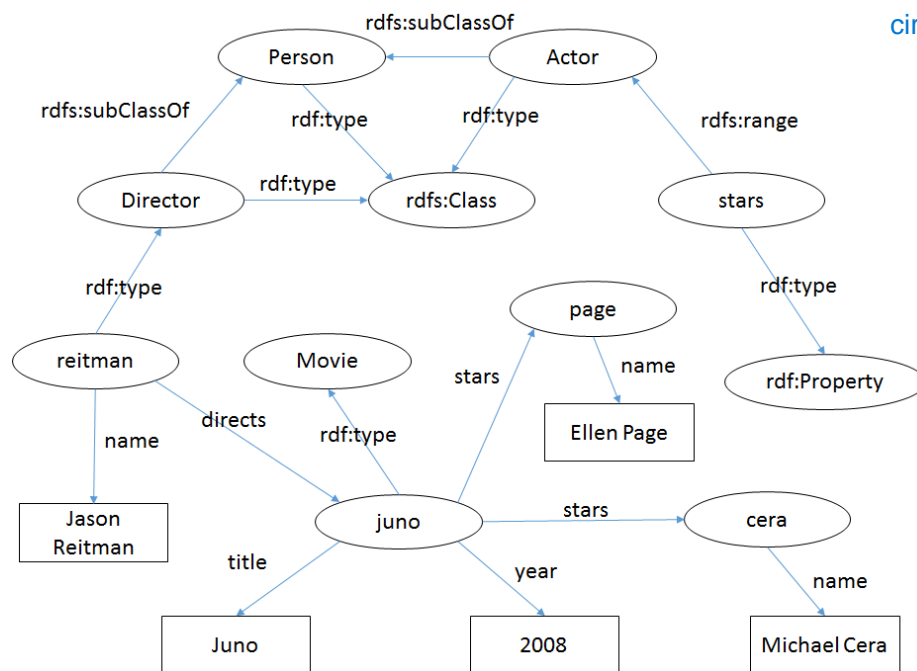
# OPEN DATA EXAM

19<sup>th</sup> of June 2015. The exam will take 2 hours. Answer each question in the provided space.  
Answers out of such space will not be considered.

Name: .....

## Question 1. [5p]

- a) Write an RDF model in Turtle notation representing the following RDF Graph. When writing the triples separate them in two blocks: one block for terminological triples (i.e., defining concepts) and one block for asserted facts (i.e., defining instances). (1,5p)



circle is concept be careful

Note: Squared concepts represent literals. For the other concepts without an explicit URI assume the <http://www.example.org/> namespace. Properly define any other prefix than `rdf` and `rdfs` you may want to use.

Prefixes:

Terminological triples (concepts)

Asserted facts (instances)

b) Assume a RDFS entailment regime for the RDF graph you just created. Represent the inferred knowledge in the form of triples and for each of them justify from what RDFS semantic construct it is inferred. Circle those inferred triples that were not already represented in the graph (i.e., clearly identify the new knowledge inferred) **(1,5p)**

c) Write a SPARQL query to retrieve the following query: *"Retrieve the number of movies starred by Ellen Page per year"*

The query must compile and include a proper definition of the prefixes used. **(1p)**

d) **Just considering this query**, what would be the best option to physically store this RDF graph: either in a relational-based or graph-based triplestore? **Justify** your answer. **(0,5p)**

not needed for this exam

e) What are the differences between a triplestore and a graph database? **(0,5p)**

not needed for this exam

## Question 2. [5p]

a) Model in RDF(S) the following statements:

*"A football player plays in a team but could have played in many teams before. Thus, we want to store the period he / she played for a team. For each team, though, he / she had a unique number assigned. Each team currently has a single manager, but obviously it might have had different managers along time. Again, we want to store the period each manager was at each team. Finally, we want to consider the figure of player-manager, who is someone that for a team played the role of manager and player at the same time".*

Draw a RDF graph modelling as many statements as you can. Clearly identify the statements modelled in your graph. Clearly identify in the graph the RDFS constructs used and the concepts created. Create your own namespace for the URIs you need to create. **(1,5p)**

b) Use Description Logics to represent those statements that you could not represent using RDF(S). Clearly identify the statement and the DL construct used. Maintain the consistency with your RDF(S) graph by using the same names (i.e., URIs) whenever possible. **(1,5p)**

c) Is there any statement that you could not represent yet? Why? **(0,5p)**

d) Now, assert the following instance to your RDF graph (draw it in the above RDF graph sketched and clearly identify it by drawing the instances within rectangles): *Messi is a player and plays for FC Barcelona from 2004 until now with the number 10.* **(0,5p)**

e) Assume now the following query over the RDF graph: *"Retrieve all managers"*. Do not consider SPARQL though and strictly **apply the open-world assumption** (answer the query from a theoretical point of view). What would be the result retrieved? Why? **(1p)**

not in syllabus