NUI Galway
OÉ Gaillimh

Semester 2 Examinations 2021/2022

Course Instance Code(s)	1CSD1, 1CSD2, 1SPE1
Exam(s)	MSc in Computer Science (Data Analytics)
Module Code(s)	CT5166
Module(s)	Knowledge Graphs
Paper No.	1
External Examiner(s)	Dr John Woodward
Internal Examiner(s)	Dr. Michael Madden *Dr. John McCrae

Instructions: Answer one question in each section. **Use a separate answer book for each section answered**.

Duration	2 hours
No. of Pages	6
Discipline(s)	Computer Science
Course Co-ordinator(s)	Dr. Frank Glavin

Requirements:

Release in Exam Venue	Yes
MCQ	No
Handout	None
Statistical/ Log Tables	None
Cambridge Tables	None
Graph Paper	None
Log Graph Paper	None
Other Materials	None
Graphic material in colour	No

Knowledge Graphs

Exam Duration: 2 Hours

Answer one question in each section

Section 1: Knowledge Graph Basics

Please answer Question 1 OR Question 2

Question 1

Question 1A **15** Marks

Consider the following paragraph

> *Tesco is a supermarket chain. Tesco sells milk, bread and cheese. 2 litres of Full-Fat Milk costs €1.69 in Tesco. There is a Tesco score located in the Galway Shopping Centre. The Galway Shopping Centre is on the Headford Road, which is in Galway. Aldi is another supermarket chain and there is an Aldi at the Westside Shopping Centre.*

Draw a knowledge graph to represent the facts contained in the text above.

Question 1B **10** Marks

Using the modelling above, write a SPARQL query that can find the supermarket that sells the cheapest milk in Galway.

Question 2

Question 2A **10** Marks

Consider the following set of triples:

```
ex:Car rdfs:subClassOf ex:Vehicle .
ex:Vehicle rdfs:subClassOf owl:Thing .
ex:Toyota ex:compactCarMake ex:Yaris .
ex:compactCarMake ex:subPropertyOf ex:make .
ex:make rdfs:range ex:Car .
```

Infer 4 triples from this dataset using RDF Schema inference.

Question 2B **10** Marks

Explain in your own words, what is meant by the open world assumption and why it is used when modelling ontologies.

Question 2C **5** Marks

Explain in your own words, what are the advantages of using URIs as persistent identifiers in RDF datasets.

Section 2: Working with Knowledge Graphs

Please answer Question 3 OR Question 4

Question 3

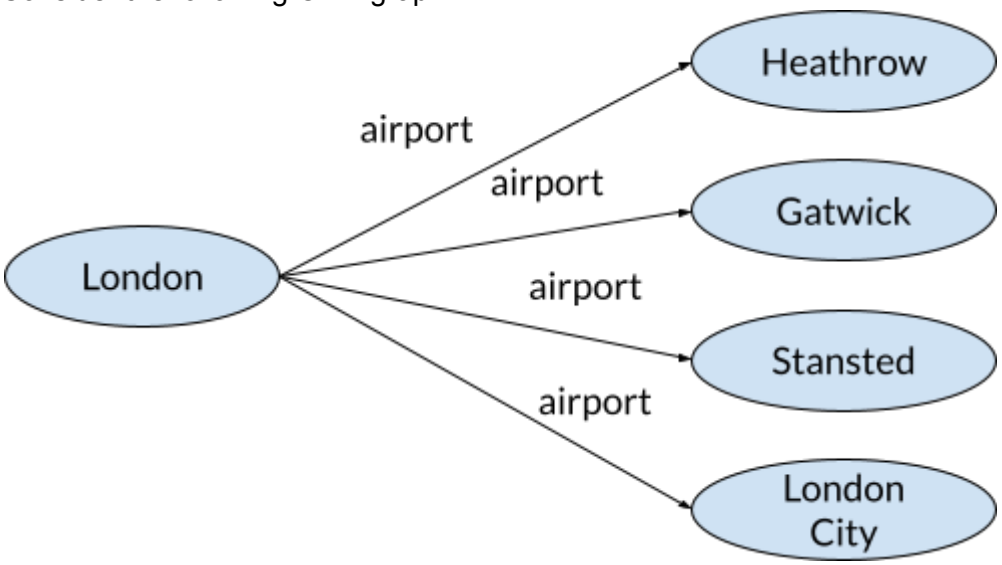
Question 3A **20** Marks

Consider the task of building an ontology about tourism, including hotels, flights and landmarks. For each of the following ontology structures give one example from the tourism domain of why you may use this ontology structure. Give a brief explanation for each:

- Inverse property
- Transitive property
- Disjoint classes
- All values from/some values from

Question 3B **5** Marks

Consider the following OWL graph.



According to this, how many airports does London actually have? Explain your answer. *(Hint: the answer is not 4).*

Question 4

Question 4A **10** Marks

What is meant by mention-level vs global relation extraction? Give one advantage or disadvantage of each approach.

Question 4B **5** Marks

Consider the following text:

When Apple Music launched, it felt clunky and limited. But it has come a long way in the past few years, with vast improvements to its functionality, music selection (75m songs), radio, playlists and other features like music videos which Spotify doesn't have.

Apple offers three months for free to new users but after that subscriptions are A$11.99 (£9.99) a month or A$119 (£99) a year, with family plans at A$17.99 (£14.99) a month. There are also bundled subscription deals with Apple's other services, such as iCloud storage and Apple TV+.

Source: The Guardian

Using Hearst patterns, extract two hypernyms from the text above.

Question 4C **10** Marks

Explain what distant supervision is and how you may use it to extract facts like the cost per month of a movie streaming service.

Section 3: Knowledge Graph Applications

Please answer Question 5 OR Question 6

Question 5

Question 5A **10** Marks

What are the 4 principles of linked data? For each of them give an explanation of how this improves the reusability of data published as linked data.

Question 5B **5** Marks

What is content negotiation? Give an example of how content negotiation works.

Question 5C **10** Marks

Data can be published as a complete archive (a ‘dump’), as linked data or through a query interface as SPARQL. For one of these methodologies, explain, with an example, why you may choose to publish data this way.

Question 6

Question 6A **15** Marks

For the enterprise data lifecycle, give an explanation with an example of a tool that could be used for each of the following steps:

- Knowledge acquisition
- Knowledge storage
- Knowledge consumption

Question 6B **10** Marks

Explain what competency questions are and how they may be used as part of the process of building an enterprise ontology.