

ICM – Computer Science Major
M1 Cyber Physical and Social Systems
Course unit on Data Interoperability and Semantics**Tutorial 3****Question 1 JSONPath**

Here is a JSON document

```
{ "store": {  
  "book": [  
    { "category": "reference",  
      "author": "Nigel Rees",  
      "title": "Sayings of the Century",  
      "price": 8.95  
    },  
    { "category": "fiction",  
      "author": "Evelyn Waugh",  
      "title": "Sword of Honour",  
      "price": 12.99  
    },  
    { "category": "fiction",  
      "author": "Herman Melville",  
      "title": "Moby Dick",  
      "isbn": "0-553-21311-3",  
      "price": 8.99  
    },  
    { "category": "fiction",  
      "author": "J. R. R. Tolkien",  
      "title": "The Lord of the Rings",  
      "isbn": "0-395-19395-8",  
      "price": 22.99  
    }  
  ],  
  "bicycle": {  
    "color": "red",  
    "price": 19.95  
  }  
}
```

Write JSONPath queries to answer the following questions:

Q1. the authors of all books in the store

Q2. the price of the third book

Q3. the last book in order.

Q4. the first two books

Q5. filter all books with isbn number

Q6. filter all books cheaper than 10

Question 2

The UNESCO Thesaurus is a controlled and structured list of terms used in subject analysis and retrieval of documents and publications in the fields of education, culture, natural sciences, social and human sciences, communication and information. Continuously enriched and updated, its multidisciplinary terminology reflects the evolution of UNESCO's programmes and activities.

Below is a screenshot of Concept <http://vocabularies.unesco.org/thesaurus/concept1670>

The screenshot displays the UNESCO Thesaurus web interface. At the top, there is a navigation bar with links for 'Vocabularies', 'About', 'Feedback', and 'Help', along with a language selector set to 'English'. Below this is a search bar with the text 'UNESCO Thesaurus' and a 'Search' button. The main content area is divided into two columns. The left column contains a list of concepts under the heading 'Alphabetical', with 'Ageing' highlighted. The right column shows the details for the concept 'Ageing', including its preferred term, broader concept, related concepts, entry terms, and translations in Arabic, French, Russian, and Spanish. The interface also includes a 'Life cycle > Ageing' breadcrumb and a 'Download this concept' section at the bottom.

Q1. which standard data models and data formats, seen in the lecture, are used to describe the UNESCO Thesaurus?

Q2. what may be the purposes of UNESCO to fund the development of a Thesaurus?

Q3. how is the UNESCO Thesaurus related to Linked Data?

Question 3

within a `<script type="application/ld+json">` tag on the html page of the tripadvisor page for the San Remo pizzeria in Saint-Étienne, one can find:

```
{
  "@context": "https://schema.org",
  "@type": "FoodEstablishment",
  "image": "https://dynamic-media-cdn.tripadvisor.com/media/photo-o/002f12/002f7f/002f80/002fd8/002fphoto0jpg.jpg?w=500&h=auto&s=1",
  "name": "Le San Remo",
  "aggregateRating": {
    "@type": "AggregateRating",
    "ratingValue": 4.5,
    "reviewCount": 350
  },
  "address": {
    "@type": "PostalAddress",
    "addressCountry": {
      "@type": "Country",
      "name": "FR"
    },
    "streetAddress": "11 rue Paul Bert",
    "addressRegion": "Auvergne-Rhône-Alpes",
    "postalCode": "42000"
  },
  "url": "https://www.tripadvisor.com/Restaurant_Review-g187269-d3179630-Reviews-Le_San_Remo-Saint_Etienne_Loire_Auvergne_Rhone_Alpes.html",
  "priceRange": "€"
}
```

Q1. What is this document ?

Q2. What are `\u002F`, `\u00F4` and `\u20AC` ?

Q2. At what URL can we find the description of the schema used here ?

Q4. At this URL, we learn that the openingHours (Text) property has the following description:

The general opening hours for a business. Opening hours can be specified as a weekly time range, starting with days, then times per day. Multiple days can be listed with commas ',' separating each day. Day or time ranges are specified using a hyphen '-'.

Days are specified using the following two-letter combinations: Mo, Tu, We, Th, Fr, Sa, Su.

Times are specified using 24:00 format. For example, 3pm is specified as 15:00, 10am as 10:00.

Write the content for this property to represent that the pizzeria is open from Tuesday to Friday 7pm-11pm, and on Saturday and Sunday 12pm-2pm and 7pm-11pm.

Question 4

The XML document below is obtained at <https://www.data.gouv.fr/fr/datasets/angers-stationnement/rdf>
It describes the dataset of all parking spots in Angers city.

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:dc:="http://www.w3.org/ns/dcat#" xmlns:dct:="http://purl.org/dc/terms/" xmlns:foaf:="http://xmlns.com/foaf/0.1/"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
  <rdf:Description rdf:about="https://www.data.gouv.fr/datasets/5e7eccbe514104d031fbf416/#resource-910a2f2e-7c5e-4cca-95ff-7df04ce61e40">
    <rdf:type rdf:resource="http://www.w3.org/ns/dcat#Distribution"/>
    <dct:identifier>910a2f2e-7c5e-4cca-95ff-7df04ce61e40</dct:identifier>
    <dct:title>angers_stationnement.json</dct:title>
    <dct:description>Angers Stationnement (json)</dct:description>
    <dc:downloadURL rdf:resource="https://data.angers.fr/api/explore/v2.1/catalog/datasets/angers_stationnement/exports/json"/>
    <dct:issued rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2020-03-26T16:19:31</dct:issued>
    <dct:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2020-03-26T16:19:31</dct:modified>
    <dct:rights>Open Data Commons Open Database License (ODbL)</dct:rights>
    <dct:license rdf:resource="http://opendatacommons.org/licenses/odbl/summary/">
    <dc:mediaType>application/json</dc:mediaType>
    <dct:format>json</dct:format>
  </rdf:Description>
  <rdf:Description rdf:about="https://www.data.gouv.fr/datasets/5e7eccbe514104d031fbf416/#resource-b50c1b8c-d9a2-4952-bf54-413fb9267163">
    <rdf:type rdf:resource="http://www.w3.org/ns/dcat#Distribution"/>
    <dct:identifier>b50c1b8c-d9a2-4952-bf54-413fb9267163</dct:identifier>
    <dct:title>Export au format GeoJSON</dct:title>
    <dct:description>- *id*: id[text]
- *nom*: nom[text]
...
- *Xlong*: xlong[text]
- *Ylat*: ylat[text]
- *tarif_1h*: tarif_1h[double]
...
  </dct:description>
  <dc:downloadURL
rdf:resource="https://data.angers.fr/explore/dataset/angers_stationnement/download?format=geojson&timezone=Europe/Berlin&use_labels_for_header=false"/>
    <dct:issued rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2020-03-28T05:04:14.375000</dct:issued>
    <dct:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2024-09-27T08:31:09.798721</dct:modified>
    <dct:rights>Open Data Commons Open Database License (ODbL)</dct:rights>
    <dct:license rdf:resource="http://opendatacommons.org/licenses/odbl/summary/">
    <dc:mediaType>application/vnd.geo+json</dc:mediaType>
    <dct:format>json</dct:format>
  </rdf:Description>
  <rdf:Description rdf:about="https://data.angers.fr/explore/dataset/angers_stationnement/">
    <rdf:type rdf:resource="http://www.w3.org/ns/dcat#Dataset"/>
    <dct:identifier>https://data.angers.fr/explore/dataset/angers_stationnement/</dct:identifier>
    <dct:title>Angers Stationnement</dct:title>
    <dct:description>Lieux de stationnement en parcs à usage public à Angers.</dct:description>
    <dct:issued rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2020-03-26T16:19:31</dct:issued>
    <dct:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2020-03-26T16:19:31</dct:modified>
    <dc:keyword>parking</dc:keyword>
    <dc:keyword>stationnement</dc:keyword>
    <dc:distribution rdf:resource="https://www.data.gouv.fr/datasets/5e7eccbe514104d031fbf416/#resource-5a979e87-6ad1-4df5-af5f-ceebee1e9424"/>
    <dc:distribution rdf:resource="https://www.data.gouv.fr/datasets/5e7eccbe514104d031fbf416/#resource-910a2f2e-7c5e-4cca-95ff-7df04ce61e40"/>
    <dc:distribution rdf:resource="https://www.data.gouv.fr/datasets/5e7eccbe514104d031fbf416/#resource-b50c1b8c-d9a2-4952-bf54-413fb9267163"/>
    ...
    <dct:publisher rdf:resource="https://www.data.gouv.fr/organizations/538346d6a3a72906c7ec5c36/">
  </rdf:Description>
  <rdf:Description rdf:about="https://www.data.gouv.fr/datasets/5e7eccbe514104d031fbf416/#resource-5a979e87-6ad1-4df5-af5f-ceebee1e9424">
    <rdf:type rdf:resource="http://www.w3.org/ns/dcat#Distribution"/>
    <dct:identifier>5a979e87-6ad1-4df5-af5f-ceebee1e9424</dct:identifier>
    <dct:title>angers_stationnement.csv</dct:title>
    <dct:description>Angers Stationnement (csv)</dct:description>
    <dc:downloadURL rdf:resource="https://data.angers.fr/api/explore/v2.1/catalog/datasets/angers_stationnement/exports/csv"/>
    <dct:issued rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2020-03-26T16:19:31</dct:issued>
    <dct:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2020-03-26T16:19:31</dct:modified>
    <dct:rights>Open Data Commons Open Database License (ODbL)</dct:rights>
    <dct:license rdf:resource="http://opendatacommons.org/licenses/odbl/summary/">
    <dc:mediaType>text/csv</dc:mediaType>
    <dct:format>csv</dct:format>
  </rdf:Description>
  ...
  <rdf:Description rdf:about="https://www.data.gouv.fr/organizations/538346d6a3a72906c7ec5c36/">
    <rdf:type rdf:resource="http://xmlns.com/foaf/0.1/Organization"/>
    <foaf:name>Angers Loire Métropole</foaf:name>
    <rdfs:label>Angers Loire Métropole</rdfs:label>
    <foaf:homepage rdf:resource="http://data.angers.fr"/>
  </rdf:Description>
</rdf:RDF>
```

Q1. Identify the main entities, their types, and the relationships that links them.

Represent the knowledge described by this document as a UML instance diagram
(only display the minimum amount of information)

Q2. The GeoJSON distribution starts like this:

```
{
  "type": "FeatureCollection",
  "features":
  [
    {
      "type": "Feature",
      "geometry": null,
      "properties":
      {
        "nb_places": 960,
        "tarif_24h": 0.6,
        "tarif_3h": 3.2,
        "abo_non_resident": 65,
        "exploitant": "ALTER",
        "nb_pr": 0,
        "moyens_acces": "AVENUE DU 11 NOVEMBRE / BOULEVARD DU MARECHAL JOFFRE",
        "gpl": "OUI",
        "tarif_pmr": "normal_payant",
        "adresse": "2 AVENUE DU ONZE NOVEMBRE 1918",
        "nb_covoit": 0,
        "nom": "Parking du Mail",
        "geo_point_2d": "47.4712547046,-0.544924904",
        "url": "http://parking-angers.fr/parking-du-mail-2/",
        "gratuit": "FAUX",
        ...
      }
    }
  ]
}
```

This document could be improved by better using GeoJSON, which is an open standard format designed for representing simple geographical features, along with their non-spatial attributes. Below is an example of a point described in GeoJSON:

```
{
  "type": "Feature",
  "geometry": {
    "type": "Point",
    "coordinates": [102.0, 0.5]
  },
  "properties": {
    "prop0": "value0"
  }
}
```

What needs to be changed in the GeoJSON dataset of Angers to better implement GeoJSON ?

Q3. Why is it problematic that the distribution doesn't properly implement the GeoJSON standard ?

Q4. On the page <https://normes.transport.data.gouv.fr/> one can read:

Public transit standards

Context

The Commission Delegated Regulation (EU) 2017/1926 of May 31, 2017 (supplementing Directive 2010/40/EU of the European Parliament and of the Council as regards the availability throughout the Union of multimodal travel information services), and in French law, LOI n° 2019-1428 du 24 décembre 2019 d'orientation des mobilités, notably in its article 25, stipulates that static and dynamic data to be opened, and applicable to NeTex and SIRI standards, must be represented by minimum national profiles, which are a declination of these standards to business needs in France.

NeTex (CEN TS 16614-1, 16614-2 and 16614-3) provides a XML-based data format and services for exchanging data describing the planned transport offer, including parkings, public transport networks, and tariffs.

An entry point to a detailed description of NeTex is at <https://normes.transport.data.gouv.fr/normes/netex/>

A NeTex description of the Angers parking spots dataset should be developed. Why, by whom, how ?

Syntax

Traversal

Array Access

Filtering

Complex Conditions

Appendix B: Overview of DCAT v2 model, showing the classes of resources that can be members of a Catalog, and the relationships between them.

