General instructions

You will be given a fixed timeframe to solve the two programming tasks. You may return your solution at any time before this deadline. Please zip your code to avoid email filters.

Programming skills

A research group is investigating what is the diagnostic yield of different genomic testing methods for renal disease. The diagnostic yield is the percentage of cases where genomic testing resulted in a confirmed diagnosis. The two main sources of data are separate software systems. One of them is capable of exporting CSV files and the second one JSON files. Examples of both files are shown below:

Json file


```
[
    {
        "patientId": "A0001",
        "differentialDiagnosis": [],
        "confimedDiagnosis": "Gitelman's syndrome",
        "test_type": "exome",
        "test_date": "2021-04-13T19:15:42.510Z"
    },
    {
        "patientId": "A0002",
        "differentialDiagnosis": [
            "Distal renal tubular acidosis",
            "Sjogren syndrome"
        ],
        "confimedDiagnosis": "",
        "test_type": "whole genome",
        "test_date": "2020-01-23T10:25:43.411Z"
    },
    {
        "patientId": "A0003",
        "differentialDiagnosis": [],
        "confimedDiagnosis": "IgA glomerulonephritis",
        "test_type": "panel",
        "test_date": "2021-01-20T12:25:43.321Z"
    }
]
```


CSV file

stud_num,pat_id,test_type,test_date,dif_diag_1,dif_diag_2,dif_diag_3,final_diag
1,897761,wgs,12/3/19,,,,gitelman syndrome
2,999871,wgs,1/12/21,bartter syndrome,gitelman syndrome,,

3,234234,exome,3/12/20,,,,familial nephrotic syndrome
4,332333,gene panel,2/2/21,,,,distal renal tubular acidosis

The FHIR standard has been selected as the common data model to represent the information required by the researchers. The following are links to the documentation of the standard:

https://hl7.org/FHIR/modules.html
https://www.hl7.org/fhir/resourcelist.html

1) Write a software program in a language of your choice that takes either of the samples and produces the corresponding representation in FHIR.

2) Design and draw an architecture diagram of a system that integrates these two sources and allows the researches to query the data. Explain how the data could be queried and give an example using a technology of your choice.