

Semantic Data Integration Report (Team 1)

Mohsen Safi Najafabadi, Mohammadamin Azin, Mohammad Almasi, Shabnam Hamzehloofard, Mona Manoochehri

Task1. Week 1 (02.05-08.05) - Datasets Selection

We collected 3 datasets about Universities this week:

- **Colleges and Universities Dataset (Mohammadamin Azin, Mona Manoochehri)**
- **National University Rankings(Mohsen Safi Najafabadi, Shabnam Hamzeloofard)**
- **University and College Campuses (Mohammad Almasi)**

All 3 dataset are uploaded to the FIM GIT LAB (CSV files)

Task1. Week 2 (09.05-16.05) - Competency Questions

- Competency Questions 1,2 (Mona Manoochehri)
- Competency Questions 3,4 (Mohammad Almasi)
- Competency Questions 5,6 (Mohammadamin Azin)
- Competency Questions 7,8 (Shabnam Hamzeloofard)
- Competency Questions 9,10 (Mohsen Safi Najafabadi)

Ten Competency Questions are Uploaded in The GIT folder of the group.

Also described CQ in Conjunctive Query form.

1- Which Universities are located in the US?

Q(Name):-

CollegesUniversities(Name, Country), Country="US"

2- What is the Url for the University of UCLA?

Q(UniversityName, Url):-

CollegeUniversityCampuses(UniversityName, Url), UniversityName="UCLA"

3- What is the contact Name for the University of ULM?

Q(ContactName):-

**CollegeUniversityCampuses(UniversityName, ContactName),
UniversityName="University of ULM"**

4- Which Universities are Located in Pulaski County?

Q(UniversityName):-

CollegeUniversityCampuses(UniversityName, County), County="Pulaski"

5- Which Universities have running > 10 and are located in NEW YORK?

Q(UniversityName):-

**CollegesUniversities(Name,...), CollegeUniversityCampuses(UnieversityName),
NationalUniversityRankings(Uni, rank, Location), Name=UnievrsityName,
UniversityName=Uni, rank>10, Location="New York"**

6- How many students are enrolled in a Bachelor degree in the universities of Harvard and Stanford?

Q(Uni, Undergrad Enrollment), Uni="Harvard University" or Uni="Stanford University"

7- Which universities are located in latitude="125217.7395" and Longitude="34.75930829"?

Q(Name)

CollegeUniversities(Name, latitude, Longitude), latitude="125217.7395", Longitude="34.75930829"

8- What is the Tuition and fees for the GeorgeTown University?

Q(Uni, Tuitionfees), NationalUniversitiesRankings(uni, tuitionfees), Uni="GeorgeTown University"

9-what is the telephone number for American College of Healthcare Sciences?

Q(Telephone), CollegeUniversities(Name, Telephone), Name="American College of Healthcare Sciences"

10- Show the number of Students and Employees at the University of Passau?

Q(StudentCount, EmployeeCount):-

CollegesUniversities(StudentCount, Name),

CollegeUniversityCampuses(EmployeeCount, UniversityName),

Name=UniversityName, Name="University of Passau"

Task1. Week 3 (17.05-23.05) - Mediated Schema

Sources/Schema/Tables

CollegesUniversities('latitude', 'Longitude', 'NAME', 'Address', 'address2', 'City', 'ZIPCode', 'Telephone', 'StudentCount', 'County', 'Country', 'OtherInfo', 'EstablishDate', 'Website', 'TotalEnrollment', 'HasDormitory', 'CapacityDormitory', 'TotalEmployee')

CollegeUniversityCampuses('UniversityName', 'Url', 'Address', 'ZIP', 'County', 'ContactName', 'EmployeeCount', 'AverageGPA', 'NumberOfGraduatedStudents', 'AnnualScholarShipGranted', 'CountStudentWorkInUni', 'NumberOfDisabledStudent')

NationalUniversitiesRankings('Uni', 'Location', 'Rank', 'Description', 'TuitionFees', 'UndergradEnrollment', 'Date', 'Country', 'ContactName', 'City', 'PostalCode')

Mediated Schema

CollegesUniversities(**NAME**, **Address** , **address2**, **City**, **ZIPCode**, **County**, **Country**, **Website**, **TotalEmployee**)

Student(**UniversityName**, AverageGPA, NumberOfGraduatedStudents, AnnualScholar,ShipGranted ,CountStudentWorkInUni, NumberOfDisabledStudent

Employee(**Uni**, **EmployeeCount**, **Location**, **Country**, **City**, **PostalCode**)

Mediated Schema is Uploaded in The GIT folder of the group.All team members contributed equally this week.

Contributions:

Mohammad Almasi:

Analyze of attribute of data set 1,2,3 to Mediate Schema.
Maintained the git repository.

Mona Manoochehri :

Analyze of attribute of data set 1,2,3 to Mediate Schema.
Maintained the git repository.

Shabnam Hamzehloofard:

Analyze of attribute of data set 1,2,3 to Mediate Schema.
Write the latex report.

Mohsen Safi Najafabadi:

Analyze of attribute of data set 1,2,3 to Mediate Schema.

Mohammadamin Azin:

Analyze of attribute of data set 1,2,3 to Mediate Schema.