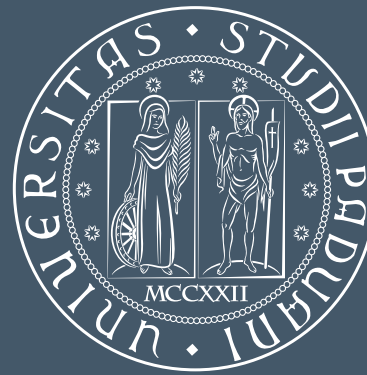


1222 • 2022  
800  
ANNI



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA

# Team Project 1: RDF



Prof. Gianmaria Silvello



[gianmaria.silvello@unipd.it](mailto:gianmaria.silvello@unipd.it)



<http://www.dei.unipd.it/~silvello/>



# Create a group

A group should be composed of three students

You need:

- A group name

- A bitbucket/GitHub repository where you upload all the required files

- A project topic: what you are modelling, and what are the data your are considering

- A URL where the data can be downloaded (e.g. Kaggle dataset, open data URL, ...)

Register your group here by Wednesday 19 October 2022

[https://docs.google.com/forms/d/e/  
1FAIpQLSeC91hdj7iHmwUgzUlz3Cq00haQLTI9LERbYsQU5\\_106a4TZQ/viewform?  
usp=pp\\_url](https://docs.google.com/forms/d/e/1FAIpQLSeC91hdj7iHmwUgzUlz3Cq00haQLTI9LERbYsQU5_106a4TZQ/viewform?usp=pp_url)

# What to do

Decide a topic/domain of interest

Find interesting data (open data is a plus)

Model the domain of interest

- Create a graph of the model (visual)

- Create a OWL file of the ontology

Python notebook with the data cleaning, ingestion and serialisation function

Provide the serialisation files of your RDF data

Ingest the data in a GraphDB instance

Provide ten insightful queries on your data

*Additional: Provide some data analytics on the data*

# Presentation

Upload all the files to the public repository you created

Everything needed to reproduce the project must be in the repository

Deadline: 21 November 2022, midnight

I'll download the files the next morning and they are the official submission

Presentation to the class: 21 and 22 December 2022

Presentation:

15 minutes + question time

Slides required: present the domain, the data, the model, the ingestion phase and the output (queries and, if available, some analytics)