COMP39/9900 Computer Science/IT Capstone Project
School of Computer Science and Engineering, UNSW

Project Number: P9

Project Title: APRA regulatory reporting automation

Project Clients: Alan Hsiao,

Project Specializations: Web application development; Software development; Cloud computing; Big data analytics and visualization; Knowledge Graph.

Background:

APRA regulatory reporting is a complex and costly compliance burden for all banks.

The process of managing regulatory requirements and updates over time is a time consuming and error prone process. The challenge is a unique one as the Australian banking regulator (APRA) must harmonise data definition across over 100 financial institutions.

Due to the complexity and inefficiency of this process, it has been a challenge for APRA to roll out additional regulatory returns, hampering its prudential governance capability as well as limiting the granularity of some of Australia's economic data made available to the Australian Bureau of Statistics. The ability to generate new economic data and metrics is crucial for economic planning purposes within the Australian government.

Number of groups: 3

Requirements and Scope:

The goal is to develop a method for managing data definitions, logic and transformations across all regulatory returns and simplify the regulatory compliance process for banks.

Requirements:

1. Develop a method to automatically catalogue APRA regulatory requirements (https://sbr-pet.apra.gov.au/ARF/ARF.html)
3. Develop a logical data model (regulatory data model) which is fit for purpose for all regulatory reports. The logical data model should be implemented as a knowledge graph (Stardog)
3. Develop a method to automate the generation of data transformation code (in SQL) from the regulatory data model to the output report formats. (using DBT)

Required Knowledge and skills:

Data modelling in both ER and graph techniques.

Web content scraping

Knowledge Graph programming and notation (RDF, TTL, Sparql)

SQL (DBT)

Cloud data lake platforms (Databricks Lakehouse)

Expected outcomes/deliverables:

A knowledge graph mapping out APRA reporting form specifications.

A web user interface to visualise the knowledge graph.

A method to specify data transformations that can be translated into executable ELT code in DBT

Supervision:

Alan Hsiao

Additional resources:

Industry partner - Cognitivo.

Cognitivo has existing IP relating to the regulatory data model, experience with graph databases and cloud data platforms.