Requesting Services from Chiang Mai Data Analytic Center

Author: Ployjun Changrue

Advisor: Assistant Professor Dr.Matinee Kiewkanya
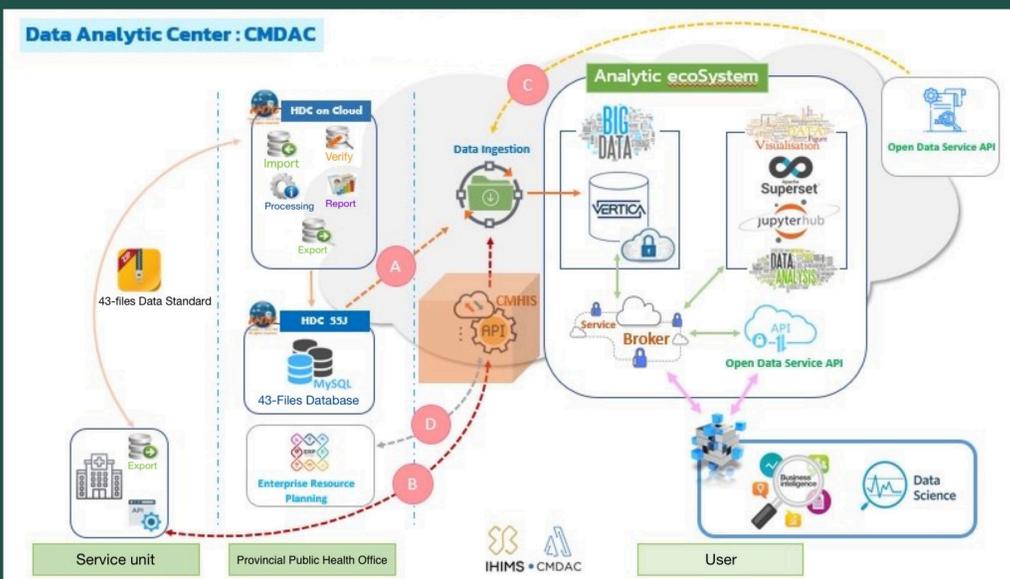
Company: Chiang Mai Provincial Public Health Office

Abstract

Data storage system in medical and public health consisted of the data of In-Patient-Department (IPD), Out-Patient-Department (OPD) and promoted activities to prevent diseases of people in Chiang Mai were collected in standard data structure form of Ministry of Public Health which were accumulated more than 10 years. In order to have efficiency analyzing of big data more than 1,500 million items, Chiang Mai Provincial Public Health Office has been developed system of "Chiang Mai Data Analytic Center, (CMDAC)". This Analytic Ecosystem consisted of analytic database system, online base of accessing and processing, data visualization and security system.

Assigned works for this co-operative education consisted of making guideline and system service accessing process for user and system testing in the role of provider and user. The details of work were collecting of data of information technology and indicating requirements of medical and public health in "Lark" documents which was able to support syntax of "Markdown" and not yet to add data in "CMDAC" system because it is in process of approval. In the part of testing as provider and user, the study was followed the Data Pipeline system of CMDAC in cancer disease of Chiang Mai. In the role of provider, "Library dbper" was used to acquire personal information and track of treat information of cancer patient in Chiang Mai from "Vertica" database by access to "Jupyterhub". After analyzing the right of "Admin" of "Superset" would be indicated to in add in data in the Superset. As the role of user, the "Dashboard" was developed in the title of "Cancer Disease in Chiang Mai" which showed information of cancer in Chaing Mai.

Introduction

The Chiang Mai Provincial Public Health Office has developed the Chiang Mai Data Analytic Center, an analytic ecosystem comprising key systems, including an analytic database, an online data access and processing system, and a data visualization system. It also features a comprehensive data security system to ensure efficient analysis of vast amounts of health data.

The assigned task primarily focused on testing CMDAC's functionality by simulating roles as both a service provider and a service recipient, particularly in analyzing cancer data in Chiang Mai. Additionally, usage guidelines and service processes were documented to support system implementation and usability.

Architecture

Results

Summary of Work Performance

- Development of Usage Guidelines: Collected IT system data and medical information requirements in Lark (Docs) using Markdown syntax. The data has not yet been added to CMDAC as it is pending approval.
- System Usage Testing: Conducted trials on CMDAC's Data Pipeline for cancer data in Chiang Mai in two roles:
  1. Service Provider: Used dbper library in JupyterHub to connect to Vertica, retrieve patient data via cancer.sql, and process it with Python. The dataset was then made accessible to the Superset Admin for integration.
  2. Service Recipient: Created a Dashboard: "Cancer Disease in Chiang Mai" in Superset by designing and organizing multiple Charts to present cancer-related insights effectively.

This work contributes to structuring CMDAC's workflow for efficient health data analysis and visualization.

Technology

Conclusion

The Chiang Mai Data Analytic Center (CMDAC) enhances public health data analysis by integrating big data processing, visualization, and security. This study focused on testing CMDAC's Data Pipeline for cancer analysis in Chiang Mai and developing usage guidelines.

By simulating service provider and user roles, the project successfully demonstrated data retrieval, processing, and dashboard creation in Superset. These contributions help improve CMDAC's usability and efficiency, supporting better data-driven decision-making in public health.

Reference

- Digital Health Group, Strategy and Planning Division, Office of the Permanent Secretary, Ministry of Public Health. (2021). Operational Manual for Data Collection and Submission Following the Standard Health Data Structure of the Ministry of Public Health (Unofficial Translation), Version 2.4, 139 pages
- Superset. Apache Superset is an open-source modern data exploration and visualization platform. from: URL: https://superset.apache.org.
- WTC. BIG DATA ANALYTIC PLATFORM SOLUTION. from: URL: https://wtc.co.th/content-new/big-data-analytic-platform-solution/.