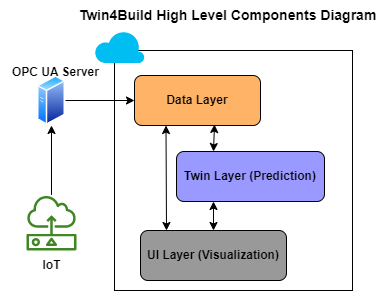
Exploring Digital Twins

Introduction

In the realm of modern architecture and facility management, the concept of digital twins has emerged as a ground-breaking innovation. A digital twin is a virtual replica of a physical building or infrastructure, enriched with real-time data and analytics, enabling a deeper understanding of its operations and performance. In this document, we delve into the intricacies of creating and utilizing digital twins specifically tailored for buildings.

Layers of Digital Twin Architecture



1. Data Layer

At the foundation of the digital twin architecture lies the Data Layer, which serves as the repository for all incoming information. For buildings, this layer aggregates data from various sources, primarily IoT sensors. These sensors include:

* Temperature sensor
* CO2 sensor
* Occupancy sensor (not applicable for SDU buildings)
* Damper position sensor
* Shade position sensor
* DMI weather data

These sensors transmit data at regular intervals, which the Data Layer processes and converts into a standardized format, such as NGSI-LD. Think of the Data Layer as a data lake, where raw information is stored before undergoing processing and analysis. The processed data is then stored in a PostgreSQL database, serving as the data warehouse for subsequent stages.

2. ML (Machine Learning) Layer or Intelligence Layer

Above the Data Layer resides the ML Layer, also known as the Intelligence Layer. This segment harnesses the power of machine learning/Physics algorithms to derive insights and make predictions based on the collected data. In the context of building digital twins, custom models tailored to specific user cases are developed by domain experts. These models are particularly vital in optimizing energy usage and mitigating CO2 emissions.

In our implementation, we have focused on HVAC (Heating, Ventilation, and Air Conditioning) systems, given their significant impact on energy consumption and environmental footprint within buildings. To operationalize these models, a robust framework and data pipeline have been established, encompassing the following steps:

1. Data Retrieval: Fetching data from the PostgreSQL data warehouse using SQL Alchemy.

2. Pre-processing: Cleaning and formatting the retrieved data for analysis.

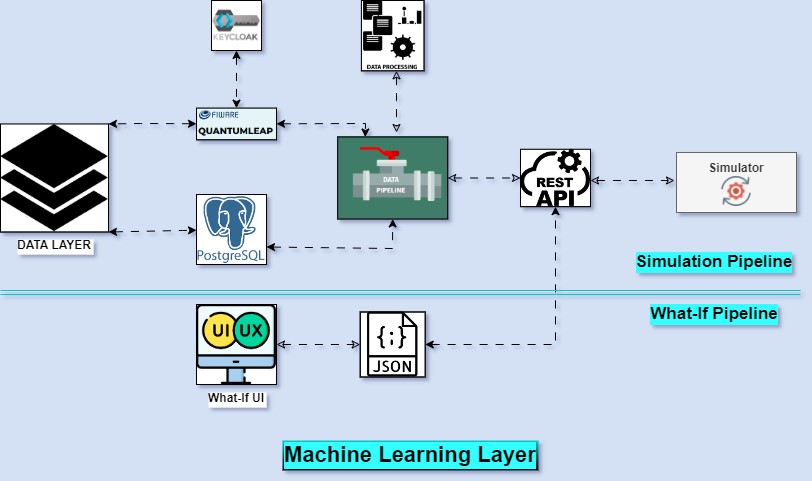
3. Validation: Ensuring the quality and integrity of the data.

4. Model Serving: Integrating the data into API endpoints for model inference.

5. Response Validation: Verifying the accuracy of model predictions.

6. Post-processing: Refining and organizing the output data.

7. Data Integration: Storing the processed data back into the database for future reference.

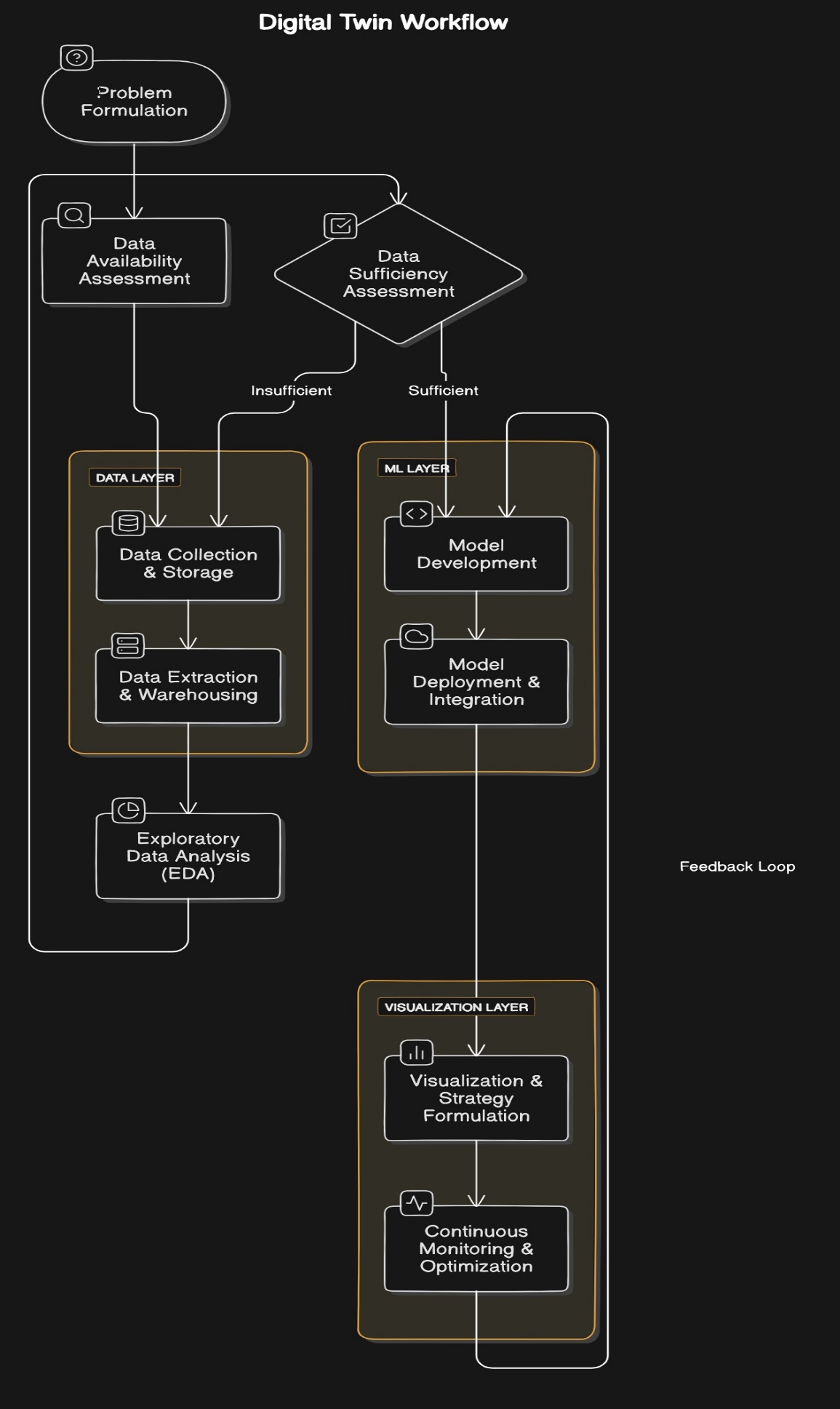


3. Visualization Layer

At the pinnacle of the digital twin architecture sits the Visualization Layer, responsible for translating complex data into actionable insights through intuitive graphical representations. Leveraging the data produced by IoT sensors and the ML Layer, this component generates custom graphs and visualizations, enabling stakeholders to monitor building performance and make informed decisions.

These visualizations empower users to analyse trends, identify anomalies, and devise strategies aimed at optimizing energy utilization and reducing CO2 emissions, thus fostering sustainability and efficiency in building operations.

**Creating a New Digital Twin: A Step-by-Step Guide**



1. **Formulate or Describe the Problem**

The journey begins with a clear articulation of the problem that the Digital Twin aims to solve. Whether it pertains to energy optimization, predictive maintenance, or occupancy management, defining the problem sets the trajectory for subsequent actions.

2. **Data Availability Assessment**

Following problem formulation, a critical inquiry ensues regarding the availability of requisite data or sensors to tackle the identified challenge. This step entails assessing whether the necessary data streams are accessible for integration into the Digital Twin ecosystem.

3. **Data Collection and Storage (Data Layer)**

Harnessing the capabilities of the existing Data Layer, the next phase involves devising strategies for data collection. Leveraging established protocols, data from diverse sources is ingested into a centralized repository, akin to a Data Lake, ensuring accessibility and scalability.

4. **Data Extraction and Warehousing (Data Layer)**

Building upon the Data Layer infrastructure, the extracted raw data undergoes transformation and enrichment processes to derive meaningful insights. Leveraging the Data Warehouse capabilities, the refined data is stored in a structured format, facilitating efficient retrieval and analysis.

5. **Exploratory Data Analysis (EDA)**

With a curated dataset at hand, the exploration phase commences, characterized by Exploratory Data Analysis (EDA). This pivotal step involves dissecting the data, identifying patterns, anomalies, and correlations, thereby laying the groundwork for informed decision-making.

6. **Data Sufficiency Assessment**

Concurrent with EDA, an evaluation is conducted to ascertain the adequacy of available data for addressing the problem delineated in the initial stage. This step serves as a litmus test to determine the feasibility and efficacy of employing machine learning or mathematical models.

7. **Model Development (ML Layer)**

Drawing upon the expertise of the ML Layer, bespoke models are crafted to tackle the identified problem statement. Whether employing machine learning algorithms or mathematical frameworks, the models are tailored to extract actionable insights from the data.

8. **Model Deployment and Integration (ML Layer)**

Upon model development, the focus shifts towards deployment and integration within the Digital Twin ecosystem. Leveraging established frameworks and APIs, the models are operationalized, facilitating real-time inference and decision support.

9. **Visualization and Strategy Formulation (Visualization Layer)**

The culmination of the process entails visualizing the results derived from the Digital Twin framework. Leveraging the Visualization Layer, custom dashboards and graphical representations are generated, empowering stakeholders to glean actionable insights and formulate strategic interventions.

10. **Continuous Monitoring and Optimization**:

Continuously monitor the Digital Twin and make adjustments as needed to optimize operations. This may involve fine-tuning system parameters, implementing energy-saving measures, or optimizing schedules.

**Case1: DT of traffic crossing, measurements on Air quality, forestation, building emissions and traffic.**

Creating a Digital Twin for a complex urban environment like a traffic crossing, incorporating measurements on air quality, forestation, building emissions, and traffic, involves a multi-faceted approach. Let's outline the step-by-step process to create this Digital Twin:

* Problem Formulation:

Begin by articulating the specific challenges and objectives of creating a Digital Twin for the traffic crossing. Identify key areas such as optimizing traffic flow, monitoring air quality, assessing environmental impact, and managing building emissions.

* Data Availability Assessment:

Conduct a comprehensive assessment to determine the availability of data sources relevant to the identified problem areas. This may include data from traffic sensors, air quality monitoring stations, satellite imagery for forestation analysis, building emission records, and traffic management systems.

* Data Collection and Storage (Data Layer):

Utilize the Data Layer infrastructure to collect and store diverse datasets from various sources. Establish protocols for data collection, ensuring real-time or near-real-time ingestion of information. This may involve integrating APIs from different data providers and IoT devices deployed at the traffic crossing.

* Data Extraction and Warehousing (Data Layer):

Implement processes to extract, transform, and store the collected data in a centralized Data Warehouse. Standardize data formats and establish data governance practices to ensure consistency and accessibility. Incorporate mechanisms for data cleansing and anomaly detection to maintain data integrity.

* Exploratory Data Analysis (EDA):

Conduct EDA on the accumulated datasets to gain insights into traffic patterns, air quality trends, forestation dynamics, building emissions, and their interrelationships. Identify correlations, outliers, and potential areas for intervention based on data-driven analysis.

* Data Sufficiency Assessment:

Evaluate the sufficiency of available data to address the objectives set forth in the problem formulation stage. Identify any gaps or limitations in the dataset and explore potential strategies to mitigate them, such as deploying additional sensors or leveraging predictive modelling techniques.

* Model Development (ML Layer):

Leverage the ML Layer to develop predictive models and analytical algorithms tailored to the specific challenges of the traffic crossing. This may involve building machine learning models for traffic flow prediction, air quality forecasting, urban greenery assessment, and emission estimation.

* Model Deployment and Integration (ML Layer):

Deploy the developed models within the Digital Twin framework, integrating them with the data processing pipeline and visualization tools. Establish APIs or microservices to enable seamless interaction between the models and other components of the Digital Twin ecosystem.

* Visualization and Strategy Formulation (Visualization Layer):

Utilize the Visualization Layer to create interactive dashboards and visualizations that communicate the insights derived from the Digital Twin. Empower stakeholders to monitor traffic conditions, air quality levels, environmental metrics, and building emissions in real-time. Formulate data-driven strategies for optimizing traffic management, mitigating pollution, promoting green initiatives, and enhancing overall urban liability.

Useful links:

[Digital\_Twin\_for\_Air\_Quality\_Management](https://www.researchgate.net/publication/372959138_Digital_Twin_DT_Smart_City_for_Air_Quality_Management)

**Case 2: DT for Public Swimming pool, measurements on water temperature, energy consumption, water turnover (circulation /filtration)**

* **Define Objectives**: Begin by outlining what you aim to achieve with the Digital Twin. This could include ensuring optimal swimming conditions, reducing energy costs, and maintaining efficient water circulation and filtration.
* **Assess Data Sources**: Identify available data sources such as sensors for water temperature, energy meters, and systems monitoring water circulation/filtration rates. Determine if additional sensors or data collection methods are needed.
* **Data Collection**: Install and configure sensors to collect real-time data on water temperature, energy consumption, and water turnover. Ensure data accuracy and reliability.
* **Data Storage and Management**: Establish a centralized data storage system to store and manage the collected data. This could be a database or cloud-based platform capable of handling large volumes of data.
* **Data Analysis**: Analyse the collected data to identify patterns, trends, and anomalies. Look for correlations between water temperature, energy consumption, and water turnover.
* **Model Development**: Develop mathematical or computational models to simulate the behaviour of the swimming pool system. These models should consider factors such as water temperature regulation, energy usage, and filtration efficiency.
* **Integration and Validation**: Integrate the developed models with the collected data to create a unified Digital Twin. Validate the accuracy and effectiveness of the models using historical data and real-world observations.
* **Visualization**: Create visualizations and dashboards to present the Digital Twin's insights in an understandable format. Visualizations may include graphs, charts, and heatmaps showing water temperature variations, energy usage trends, and circulation/filtration rates.
* **Continuous Monitoring and Optimization**: Continuously monitor the Digital Twin and make adjustments as needed to optimize swimming pool operations. This may involve fine-tuning system parameters, implementing energy-saving measures, or optimizing filtration schedules.
* **Decision Support**: Utilize the Digital Twin to make data-driven decisions regarding swimming pool management. Use insights from the Digital Twin to improve water quality, enhance energy efficiency, and ensure a pleasant swimming experience for users.

**Useful Links digital twin for pools**:

[**Digital Twin for swimming pool developed by EDAG**](https://insights.edag.com/en/smart-city-reducing-swimming-pool-costs-simulation-digital-twin)

[**Digital Twin For swimming pools trends**](https://www.aquatechtrade.com/news/utilities/digital-twins-in-water)