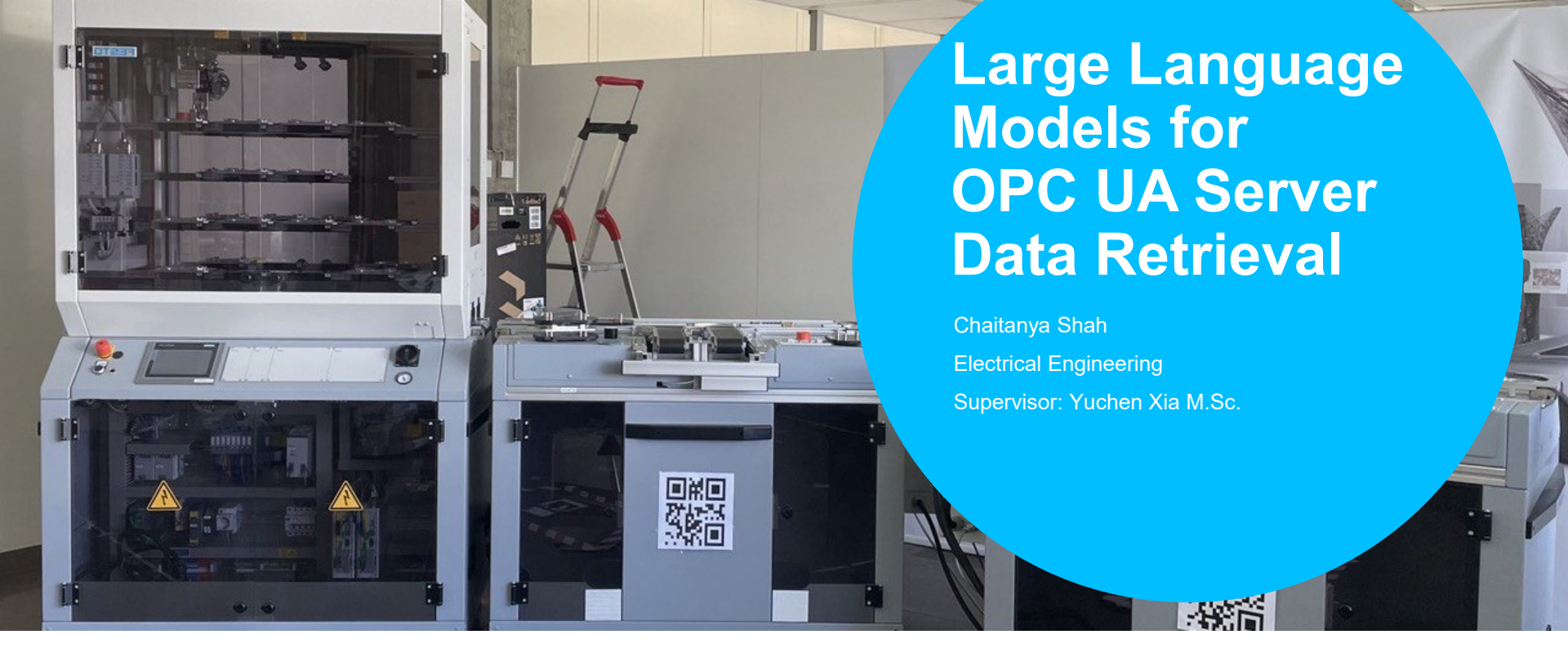
University of Stuttgart
Institute of Industrial Automation
and Software Engineering



Large Language Models for OPC UA Server Data Retrieval

Chaitanya Shah
Electrical Engineering
Supervisor: Yuchen Xia M.Sc.

Agenda

- Result Preview
- Motivation
- Basis
- Conceptual Design
- Implementation
- Evaluation
- Summary and Outlook
- Q&A Session

Motivation

Motivation:

Group of people looking for solution of specific problem by going through massive stack of data.

Trainee engineer finding the same problem solution by using LLM enabled data retrieval tool designed for the field of industrial automation.

Motivation:

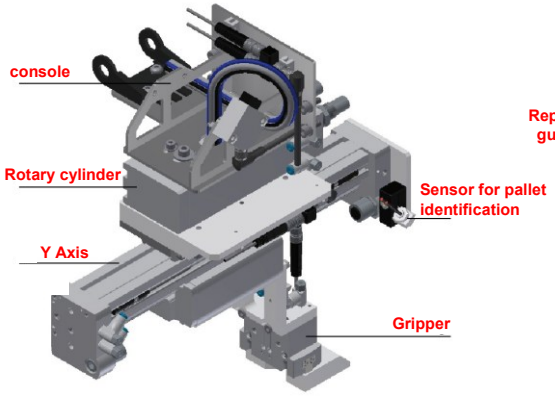
Benefits:

- Ease of managing Data Complexity and Volume
- Improved Data Integrity
- Increased Operational Efficiency and Productivity
- Streamlined Maintenance and Troubleshooting
- Optimized Resource Utilization

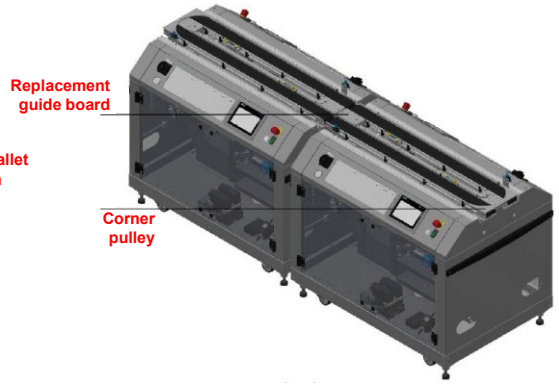
Basis

Basis: CP Factory

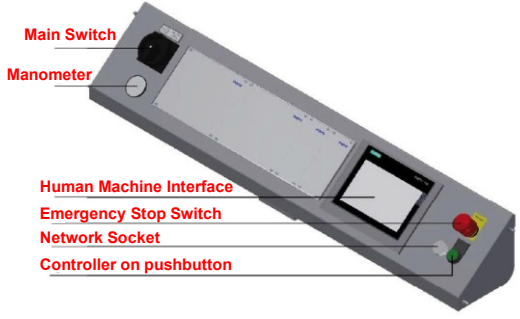
CP Factory: ASRS for pallets



Gripper Assembly of ASRS for pallets

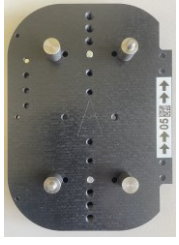


Conveyor Assembly of ASRS for pallets



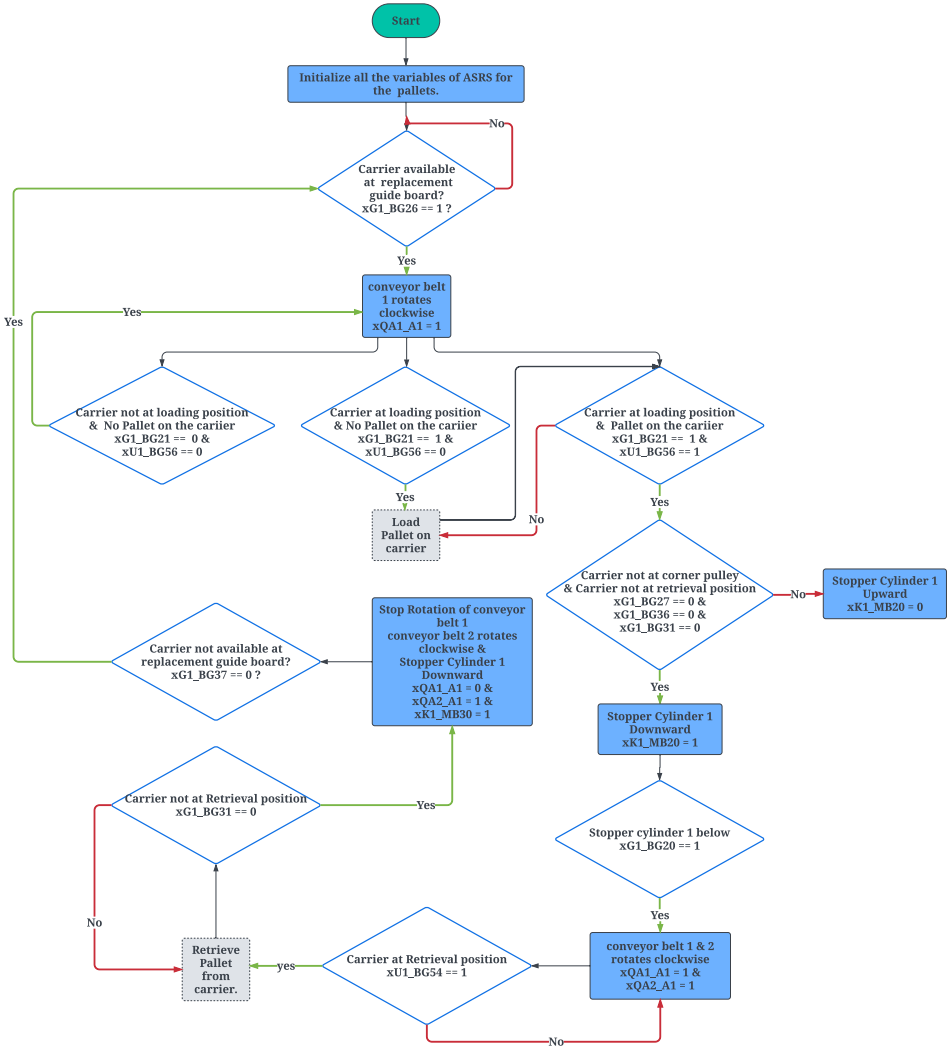
The Control Unit of ASRS for pallets

Pallet



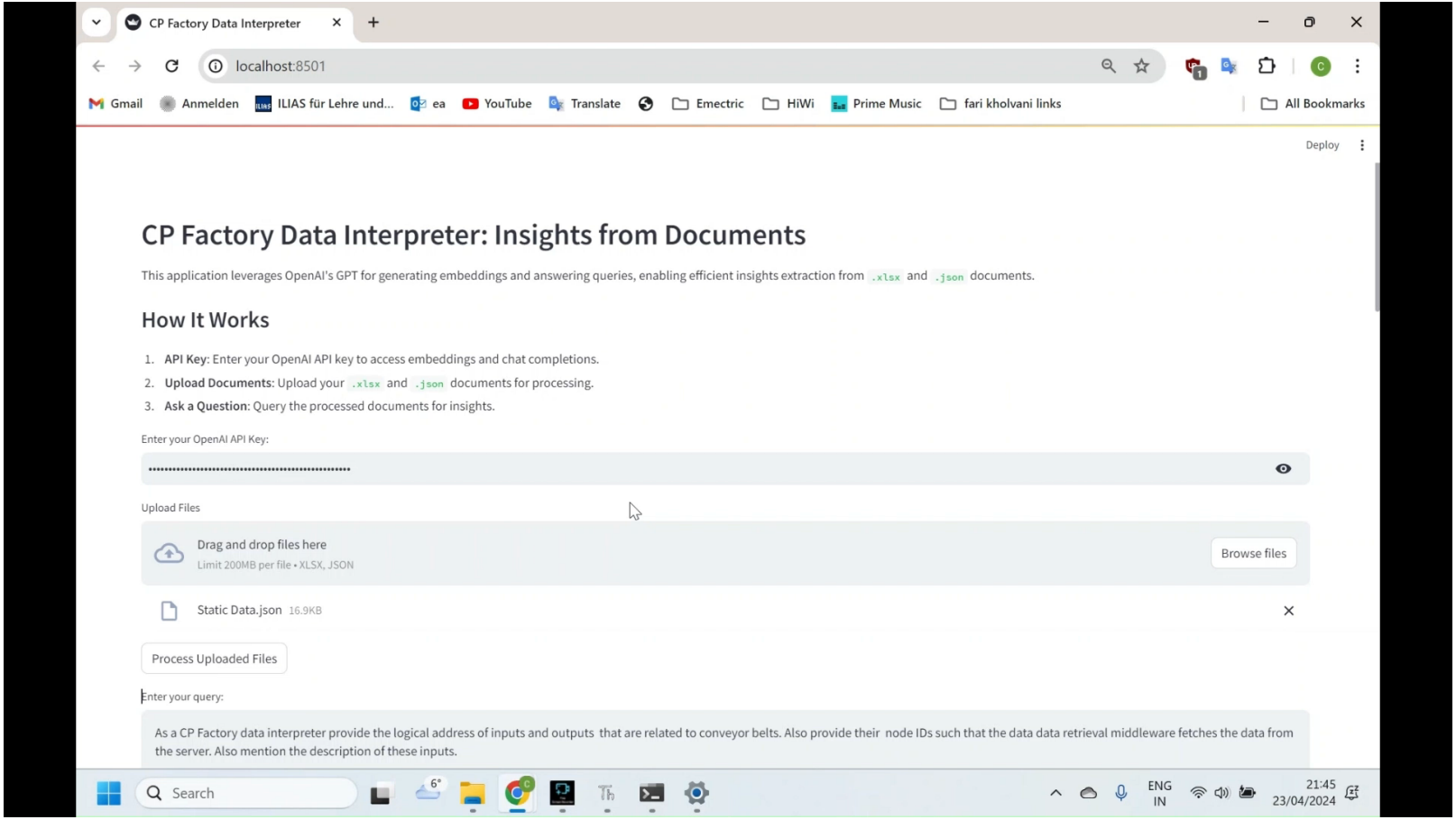
Carrier

Basis: CP Factory



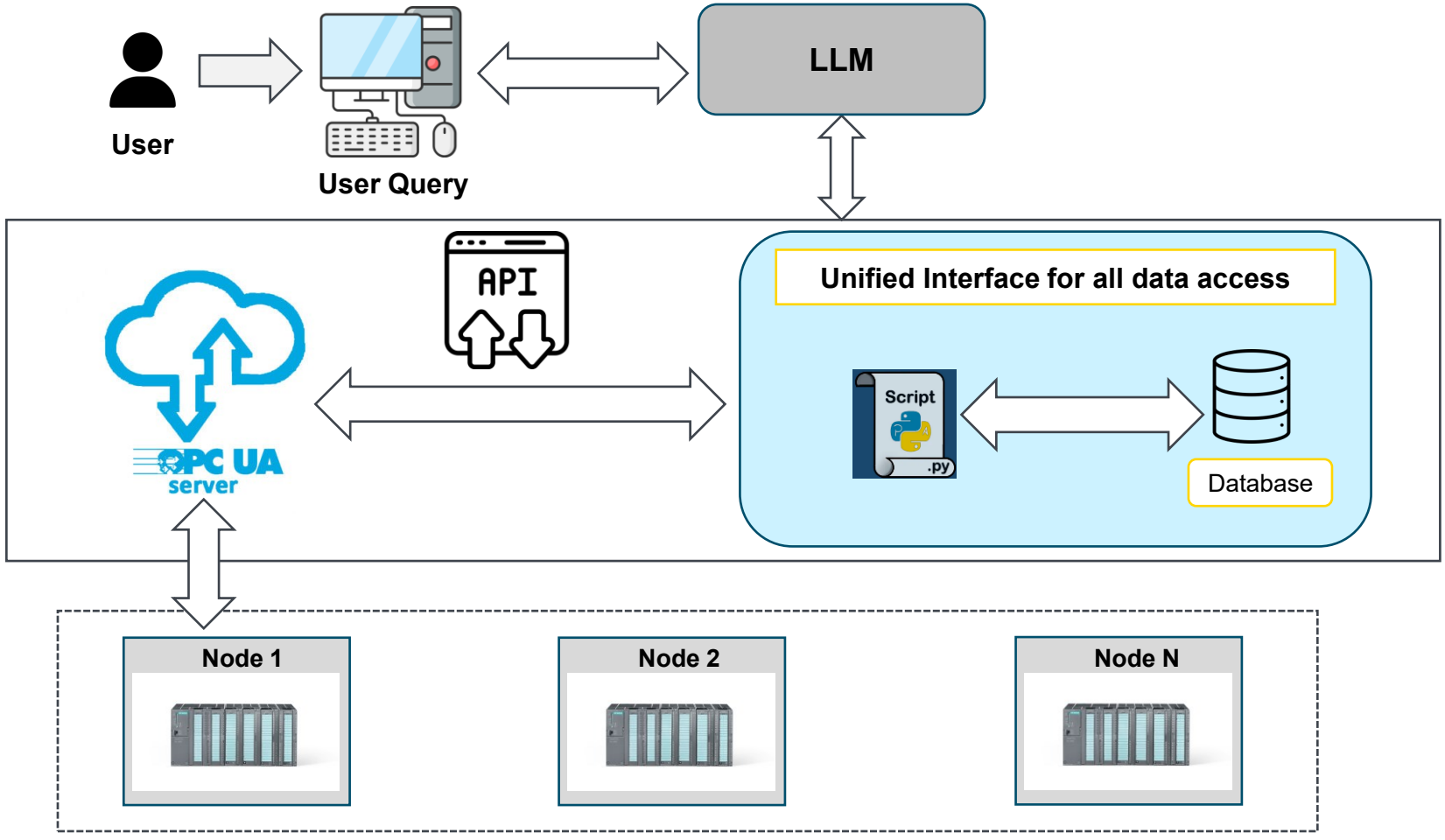
Result Preview

Result Preview:



Conceptual Design

Conceptual Design of LLMs for OPC UA Server Data Retrieval:



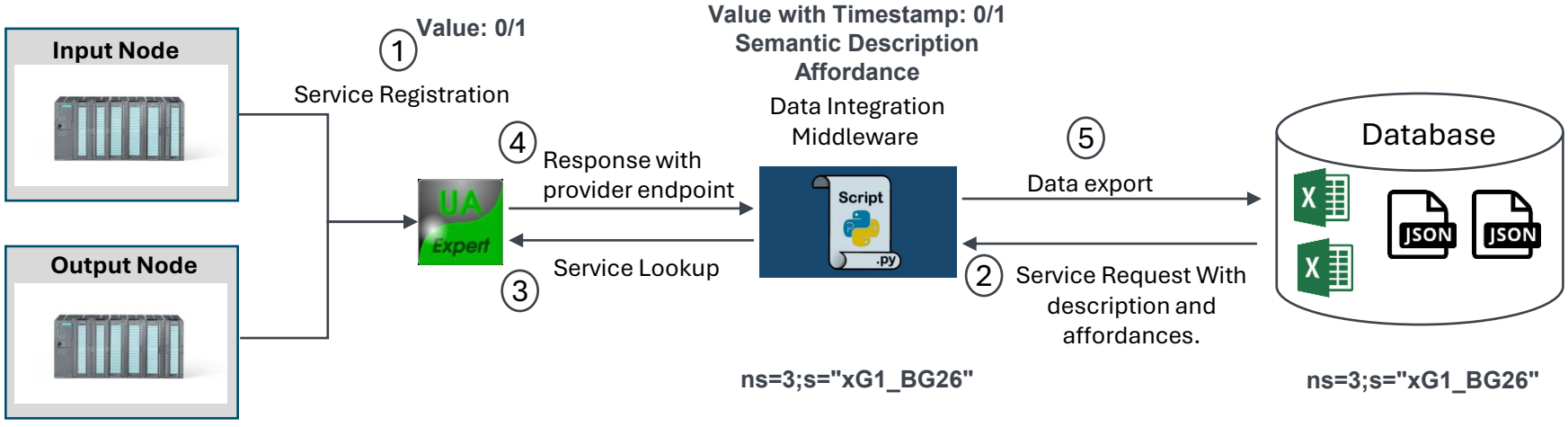
Implementation

- Dataset Creation
  1. Static Data (Metadata)
  2. Dynamic Data (Semantic Description)
- Response Generation using LLM
- User Interface

Implementation: Static Data

Field	Description
Node ID	Endpoint of the OPC UA.
Semantic Name	Name of the node mentioned on the OPC UA server.
Data Type	Type of data such as Boolean, Integers, Float, etc.
Position	Placement of the sensors or actuators in CP Factory.
Logical Address	Physical address of the PLC where the sensors and actuators are connected
Description	Functionality of the CP Factory
Notes	• Sensor Model Number (Input) • Location of Software's Function Block and Rung Numbers (Output)
Source	Literature from which the metadata for the CP Factory is extracted • Inputs: The user manual of the ASRS for the pallets provided by Festo Didactic. • Output: The information is interpreted from the system requirements and the program (software) on which CP Factory's ASRS for the pallet system is working.

Implementation: Fetching Dynamic Data from OPC UA Server for Database Creation



Field	Description
Value at the timestamp	The value corresponds to the actual endpoint data with timestamp from the OPC UA.
Semantic Description and Affordances	Description written over following question for state change from “**true-false**” and “**false-true**”. • What is the Fact? • What should happen if any of the inputs or output changes their states? • What is the reason behind the change of any inputs or outputs or simply why it happened?

JSON Files Example for Static and Dynamic Data

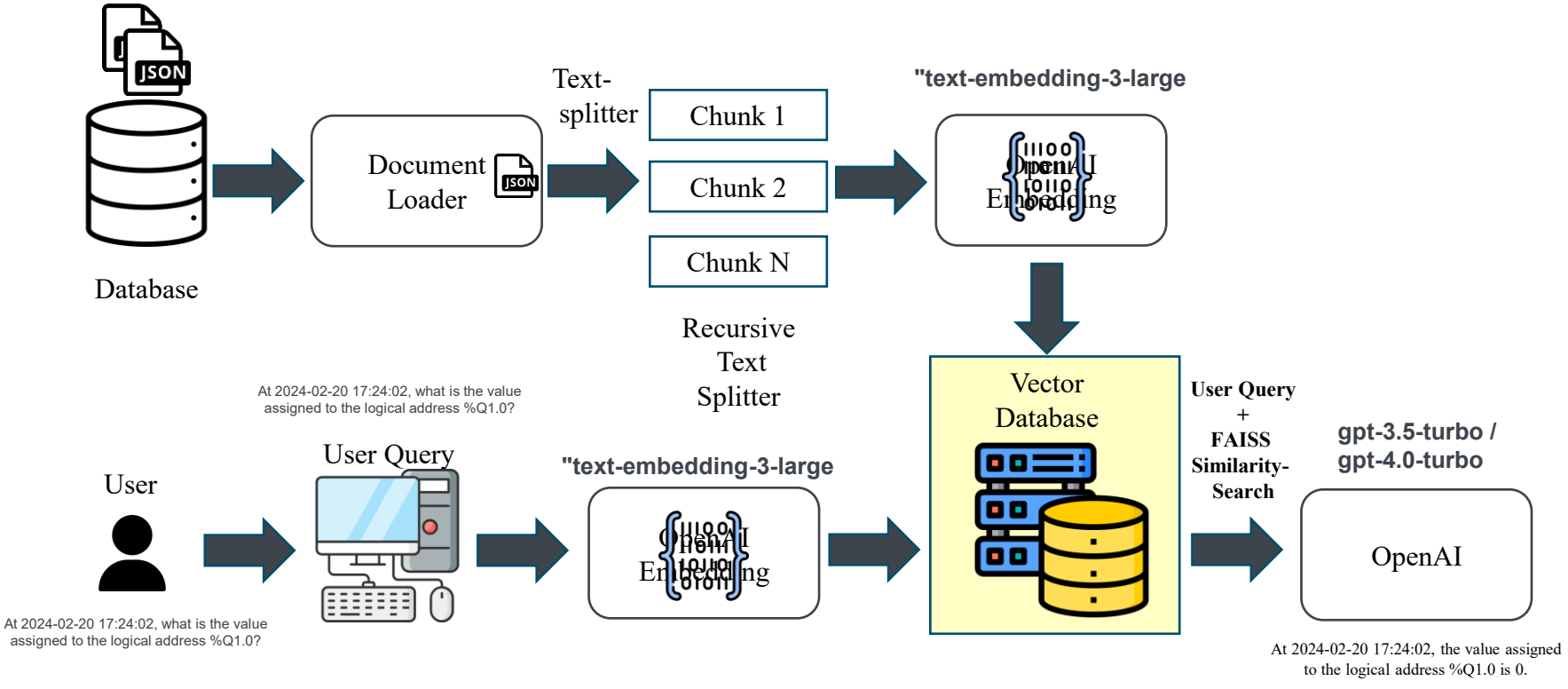
Static Data

```
{
    "Input":[
     {
      "Node_ID": "ns=3;s=\"x24VNA\"'",
      "Position": "User Interface Panel of ASRS for pallets.",
      "Semantic Name": "x24VNA",
      "Logical Address": "%I0.0",
      "Description": "0=Emergency stop button pressed",
      "Notes": "Emergency stop button non-illumin(M22-PV).",
      "Source": "1) \"04.1 CP Factory_LA (1512)V03\" 2)
      \"CP-F-ASRS32-P-GB-A001\" 3)\"CP-F-ASRS32-P-GB-C001\""
     }
    ],
    "Output":[
     {
      "Node_ID": "ns=3;s=\"xQA1_A1\"'",
      "Position": "Conveyor Belt Assembly of ASRS for
      pallets",
      "Semantic Name": "xQA1_A1",
      "Logical Address": "%Q0.0",
      "Description": "1=belt drive1 activate clockwise
      rotation",
      "Notes": "You can find logic in Actuator control > 4Q
      Drive with error monitoring > DriveMon_4Q > Network 16,
      17, 21, and 58.",
      "Source": "CP-F-1590110-DE-Stuttgart_20201002_V15.1.
      ap15_1"
     }
    ]
}
```

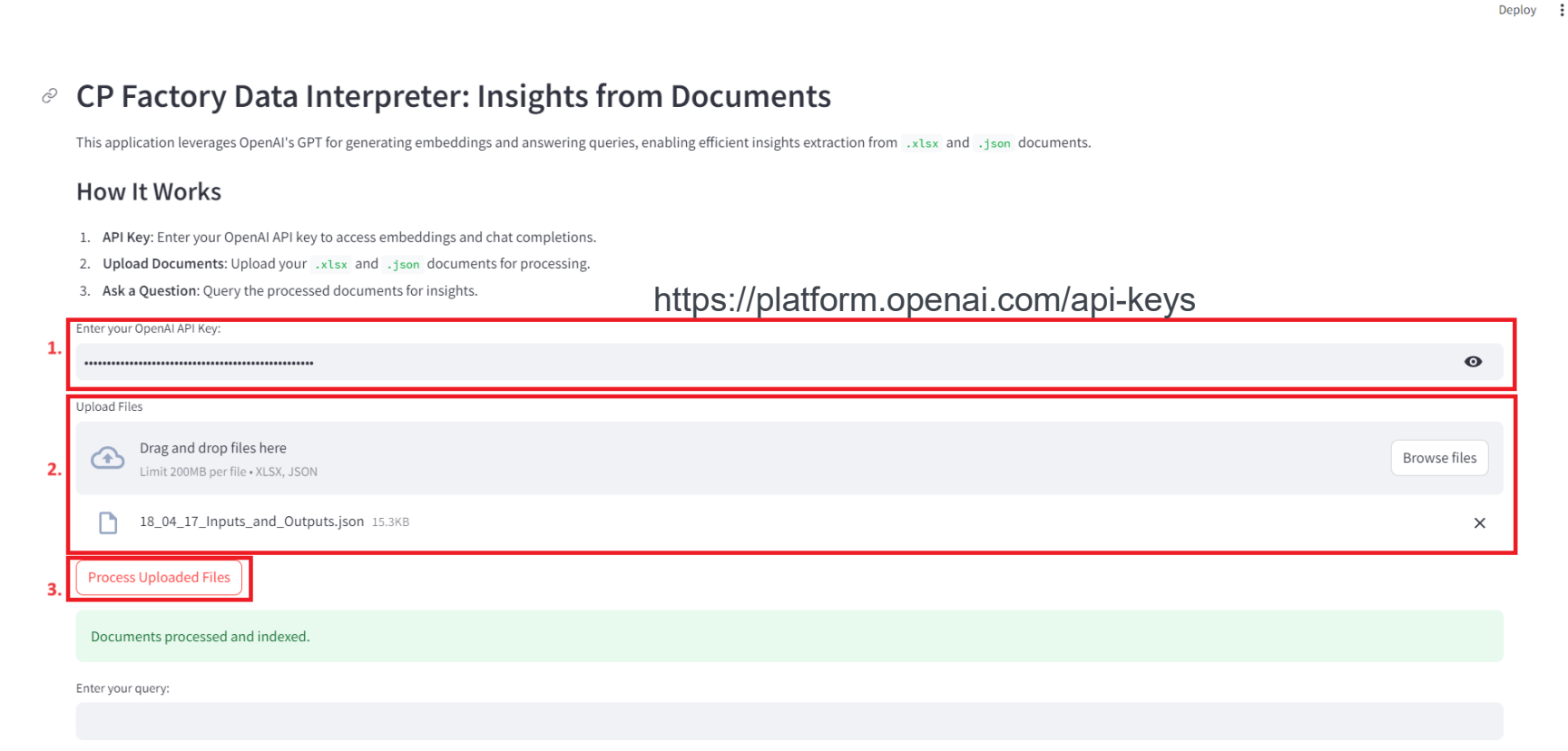
Dynamic Data

```
{
    "df1": [
        {
            "Semantic Name": "xG1_BG26",
            "Logical Address": "%I1.6",
            "Description": "Conveyor Belt 1 Infeed",
            "Detailed Description": "The carrier is at the
            replacement guide board. It should go to loading
            position of the CP storage. This behavior change
            of state denotes that a carrier is fed through CP
            branch.",
            "Value in Binary @2024-02-20 18:04:27.590593": 0,
            "Value in Binary @2024-02-20 18:04:28.604061": 1
        }
    ],
    "df2": [
        {
            "Semantic Name": "xQA1_A1",
            "Logical Address": "%Q0.0",
            "Description": "1=belt drive1 activate clockwise
            rotation",
            "Detailed Description": "The conveyor belt 1
            should drive in clockwse direction. The change in
            behavior indicates that either the carrier is
            available at the replacement guide board of the
            conveyor belt 1 or there is a input \"Right\" for
            conveyor belt 1 from the HMI.",
            "Value in Binary @2024-02-20 18:04:27.881609": 0,
            "Value in Binary @2024-02-20 18:04:28.901734": 1
        }
    ]
}
```

Implementation: Response Generation using LLM

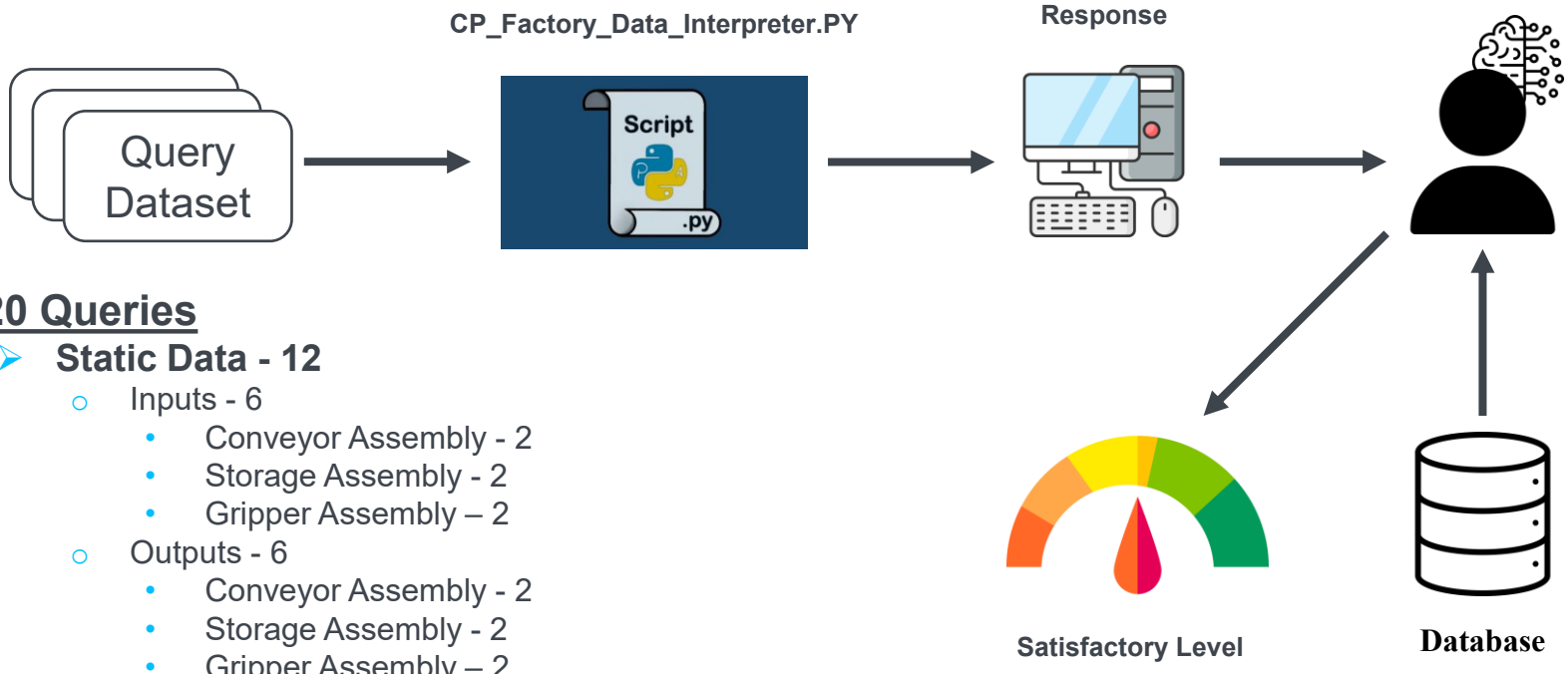


User Interface:



Evaluation

Quantitative Assessment Test for Concept evaluation:



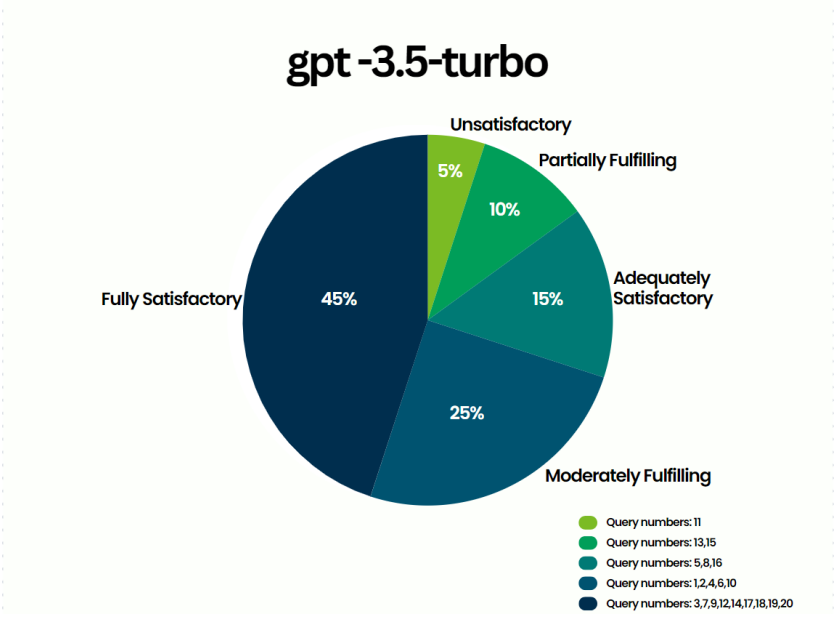
- **20 Queries**
  - **Static Data - 12**
    - Inputs - 6
      - Conveyor Assembly - 2
      - Storage Assembly - 2
      - Gripper Assembly – 2
    - Outputs - 6
      - Conveyor Assembly - 2
      - Storage Assembly - 2
      - Gripper Assembly – 2
  - **Dynamic Data - 8**
    - Inputs - 4
      - Conveyor Assembly – 2
      - Gripper Assembly – 2
    - Outputs - 4
      - Conveyor Assembly – 2
      - Gripper Assembly – 2

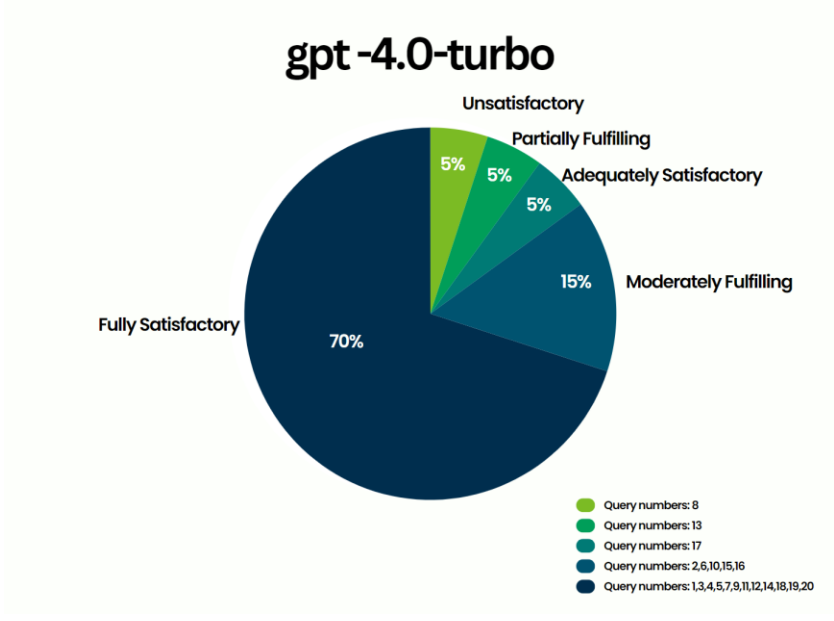
1 - Unsatisfactory
2 - Partially Fulfilling
3 - Adequately Satisfactory
4 - Moderately Fulfilling
5 - Fully Satisfactory

Weighing Criteria for response generated by GPT Models:

Weight	Comment	Reasoning
1	Unsatisfactory	The response generated by the GPT model is completely irrelevant or incorrect or fails to address the user's query in any meaningful way.
2	Partially Fulfilling	The response addresses the user's query but in a limited or superficial manner. It may omit key data or essential details.
3	Adequately Satisfactory	The response correctly addresses the query with insignificant errors which provides the basic correct answer but may offer vague or incomplete insights.
4	Moderately Fulfilling	The response is thorough delivering relevant details and describes a good understanding of the topic. However, it may either slightly miss the accuracy needed or include excessive information that could potentially confuse the user.
5	Fully Satisfactory	The response fully satisfies the query with comprehensive details, and deep insights, and caters concise response that addresses the query thoroughly.

Evaluation Result:





	gpt-3.5-turbo	gpt-4.0-turbo
Total Weights out of 100	79	86
Result of Test	85%	90%

Summary & Outlook

Summary & Outlook

Summary:

- Successfully integrated LLM capabilities to extract OPC UA server data's semantics description and affordances.
- Enhanced user experience in interacting with complex cyber-physical systems.

Outlook:

- Explore automating the addition of nodes and writing semantic information using the LLM model.
- Potential for real-time interaction with the system.
- Explore expanding the concept to control the cyber-physical factory directly.

University of Stuttgart
Institut of Industrial Automation
and Software Engineering

Thank you!

Chaitanya Pareshkumar Shah

e-mail st181599@stud.uni-stuttgart.de
phone +49 (0) 711 685-
fax +49 (0) 711 685-

University of Stuttgart
Institute of Automation and software systems
Pfaffenwaldring 47, 70569 Stuttgart