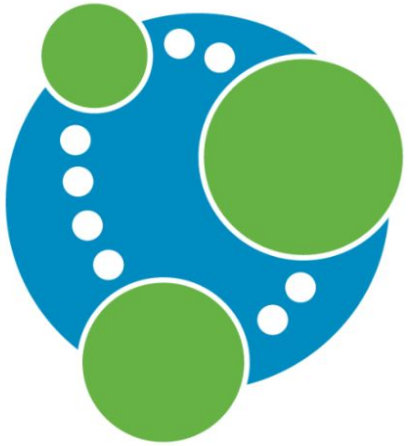
BPMN to Neo4j

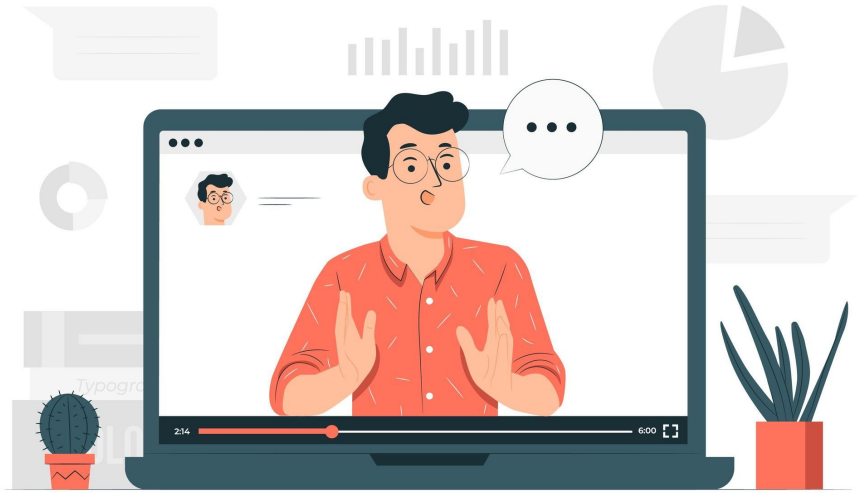


Mentor: Prof. Mridul Sankar Barik

By

- *Kushal Das 002010501071*
- *Rohit Sadhu 002010501074*
- *Soumyadeb Misra 002010501088*
- *Arnab Sur 002010501034*

Introduction



What is BPMN ?

- *BPMN stands for Business Process Model and Notation.*
- *Standardized method for understanding business processes by depicting them in a visual manner*
- *Consist of a set of standardized symbols to represent different elements of a business process*

Example BPMN Diagram

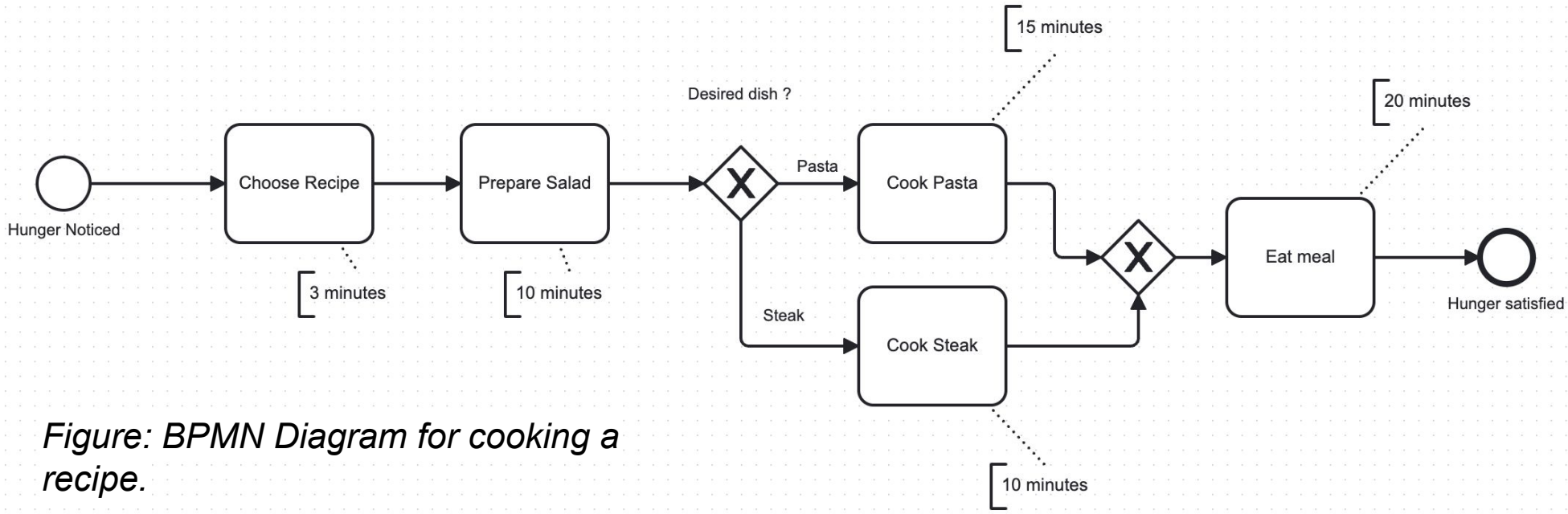


Figure: BPMN Diagram for cooking a recipe.

Background and Motivation

- **Why BPMN?**
  - Can be directly translated into executable process models
  - Primary goal to support business process management
- **Why graph database?**
  - In RDBMS, join queries performance reduces when datasets are larger which remain constant in a graph database.
  - Queries are localised to a portion of the graph.
  - In present day most information is presented in form of a graph like social network, web page graphs etc.

Survey of relevant works

- **A BPMN-based language for modeling corporate communications**
  - The introduction of CCML for visualizing organizational communication with BPMN through structured language extensions.
- **Privacy-enhanced BPMN: enabling data privacy analysis in business processes models**
  - The paper introduces PE-BPMN for privacy-enhanced process modeling, addressing GDPR concerns. It proposes a stratified model, taxonomy, syntax, and analysis techniques.
- **A Framework for Querying Graph–Based Business Process Models**
  - It propose BPMN-Q, a visual query language, for querying and reusing graph-based business process models efficiently, using relational databases for processing.

Problem Statement

Design an application that allows seamless integration between **BPMN.js** diagrams and **Neo4j** graph databases to facilitate efficient visualization and management of business processes.

Proposed Methodology

Involved steps:

- **XML Representation:**
  - BPMN diagrams serialized into XML format using the BPMN XML Schema. This XML representation captures all the necessary information.

- **Extract BPMN Information:**
  - BPMN elements need to be extracted and analyzed to understand the structure and flow of the process.
  - Elements include Tasks, Gateways, Start/End Events, Data Store etc.

Proposed Methodology (Contd.)

- **Mapping to Graph Model:**
  - Mapping the extracted information to graph model to represent in the form of Neo4j nodes & relationships.

- **Node and Relationship Mapping:**
  - **Nodes**: Each BPMN element (task, gateway, event) can be represented as a corresponding node in Neo4j.
  - **Relationships**: The connections between BPMN elements (sequence flows) can be represented as relationships in Neo4j.

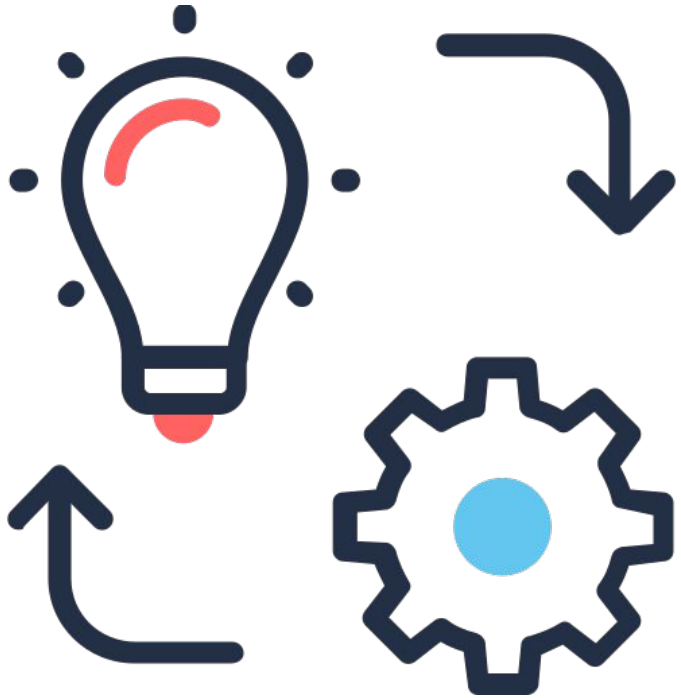
Proposed Methodology (Contd.)

- **Property Mapping:**
  - Attributes of BPMN elements can be mapped to properties of nodes and relationships in Neo4j.
  - For example, annotations, names etc. of nodes.
- **Data Import:**
  - Importing the above data into Neo4j.
  - Neo4j provides tools and APIs for importing data from various formats, including XML.
- **Querying and Analysis:**
  - Querying the data using Cypher (Neo4j's query language) to analyze the process.

Sample BPMN XML

```
<?xml version="1.0" ...>
  <bpmn2:process id="Process_1">
    <bpmn2:startEvent id="Event_0d5lr0o" name="Hunger Noticed">
      <bpmn2:outgoing>Flow_0c3sy3l</bpmn2:outgoing>
    </bpmn2:startEvent>
    <bpmn2:task id="Activity_1p8iqcn" name="Choose Recipe">
      <bpmn2:incoming>Flow_0c3sy3l</bpmn2:incoming>
      ...
    </bpmn2:task>
      <bpmn2:textAnnotation id="TextAnnotation_1q695hs">
      <bpmn2:text>3 minutes</bpmn2:text>
    </bpmn2:textAnnotation>
    <bpmn2:association id="Association_041a4y2"
sourceRef="Activity_1p8iqcn" targetRef="TextAnnotation_1q695hs" />
    <bpmn2:sequenceFlow id="Flow_0c3sy3l" sourceRef="Event_0d5lr0o"
targetRef="Activity_1p8iqcn" />
    <bpmn2:exclusiveGateway id="Gateway_1oizb5t">
      ...
    </bpmn2:exclusiveGateway>
    <bpmn2:endEvent id="Event_0vdzu5q" name="Hunger satisfied">
    </bpmn2:endEvent>
  </bpmn2:process>
</bpmn2:definitions>
```

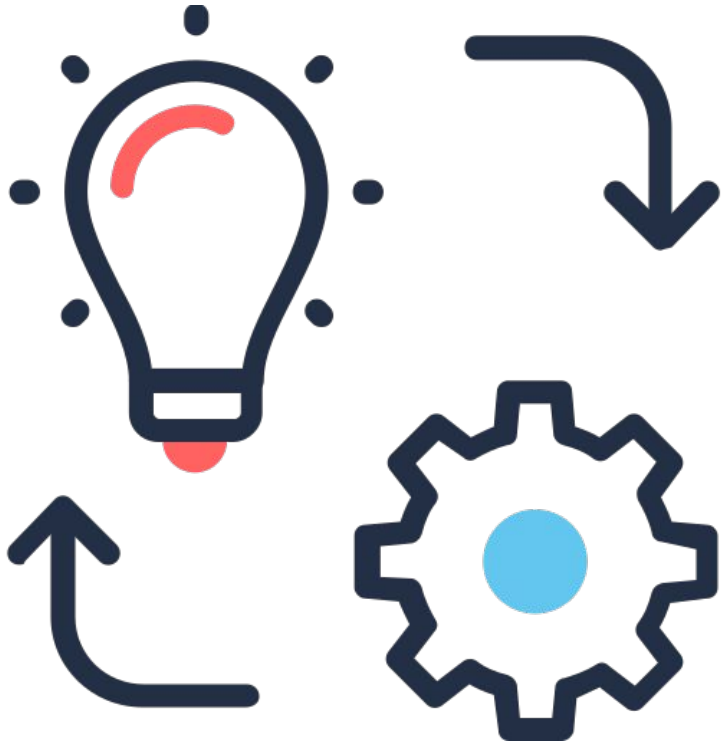
Implementation



- **XML Representation:**
  - Using BPMN moddle to serialise the BPMN diagram to XML
  - BPMN Moddle - Read and write BPMN 2.0 diagram files in NodeJS and the browser

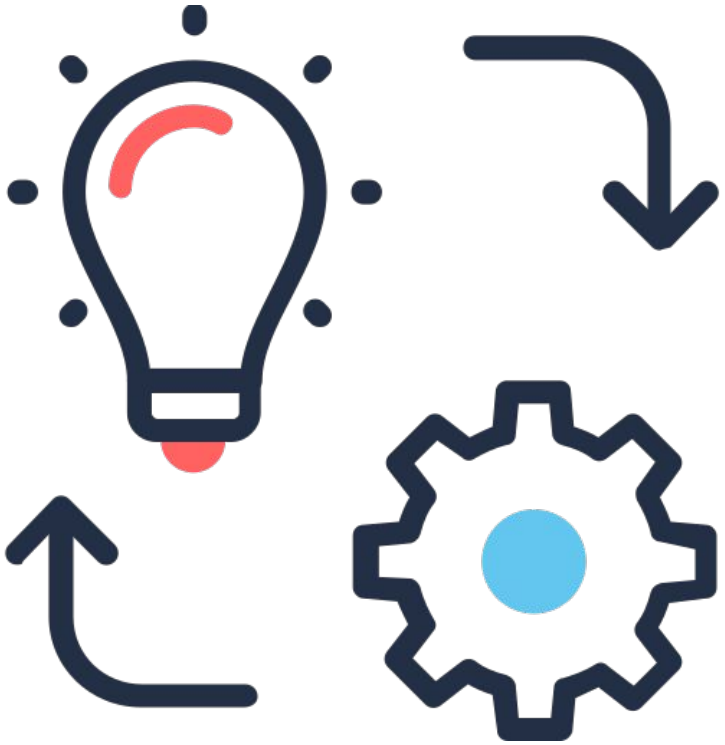
- **Extract BPMN Information:**
  - Converting BPMN XML to Javascript object using xml2js npm package
  - Creates a hierarchical javascript object
  - Supports callback and promise handling.

Implementation (contd.)



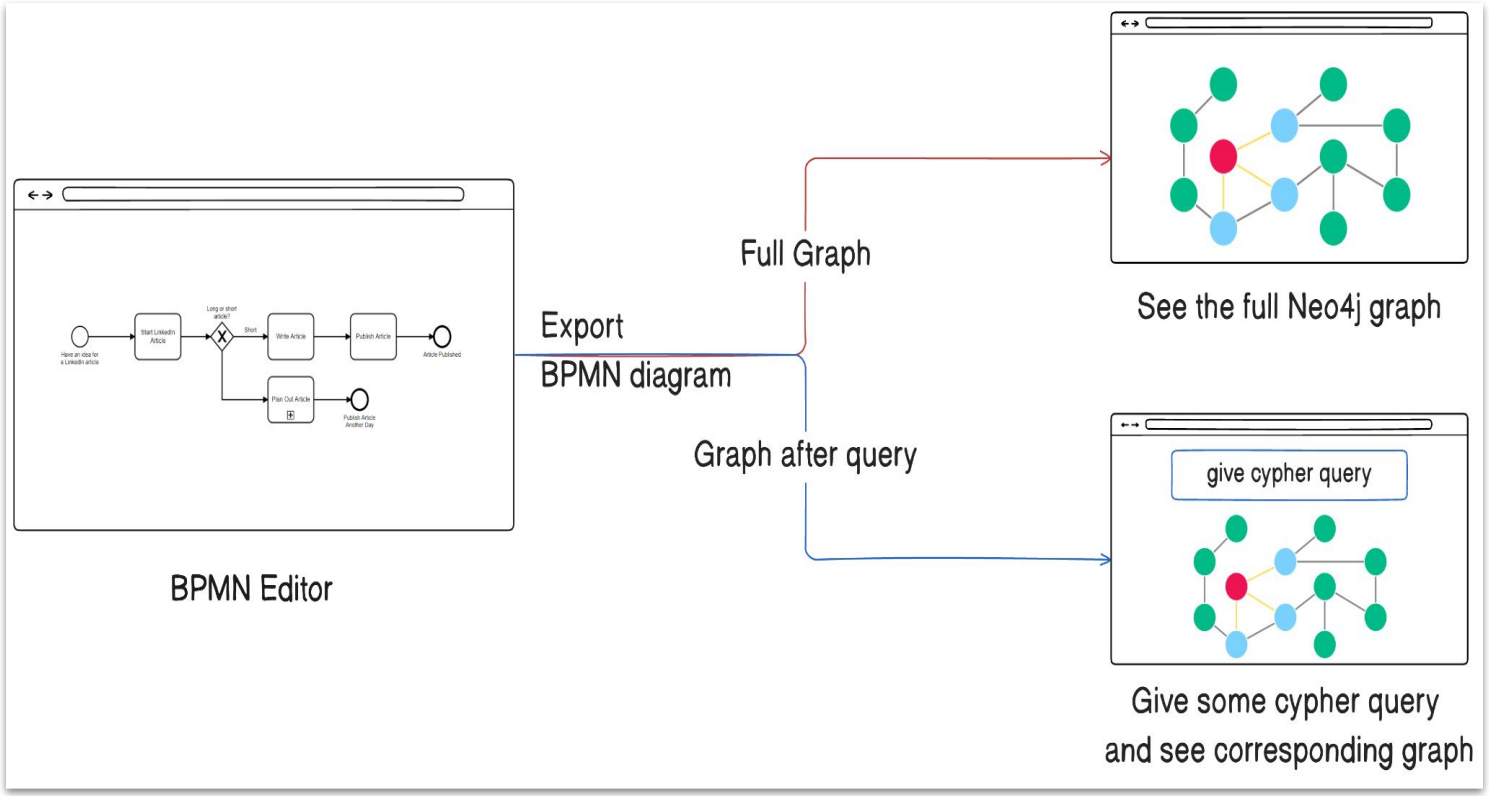
- **Mapping:**
  - **Nodes**: Extracting the tasks from javascript object to nodes array.
  - **Relationships**: Extracting the relationships from the sequence flows in the javascript object to relationships array.
- **Property Mapping:**
  - Annotations and properties are added to the nodes objects in the nodes array

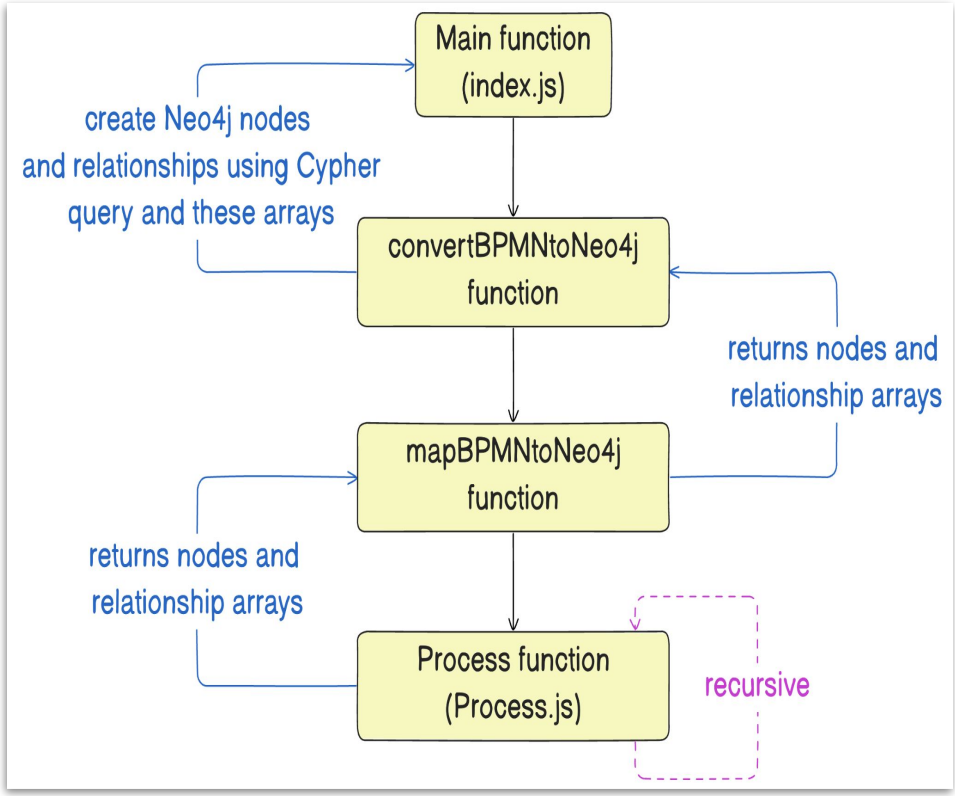
Implementation (contd.)



- **Data Import:**
  - Neo4j driver running cypher query to populate the Neo4j database.
  - Neo4j credentials taken from env variables from frontend.
- **Querying and Analysis:**
  - Neo4j driver instance in the frontend for querying and analysing the existing graph database.

Program Flow





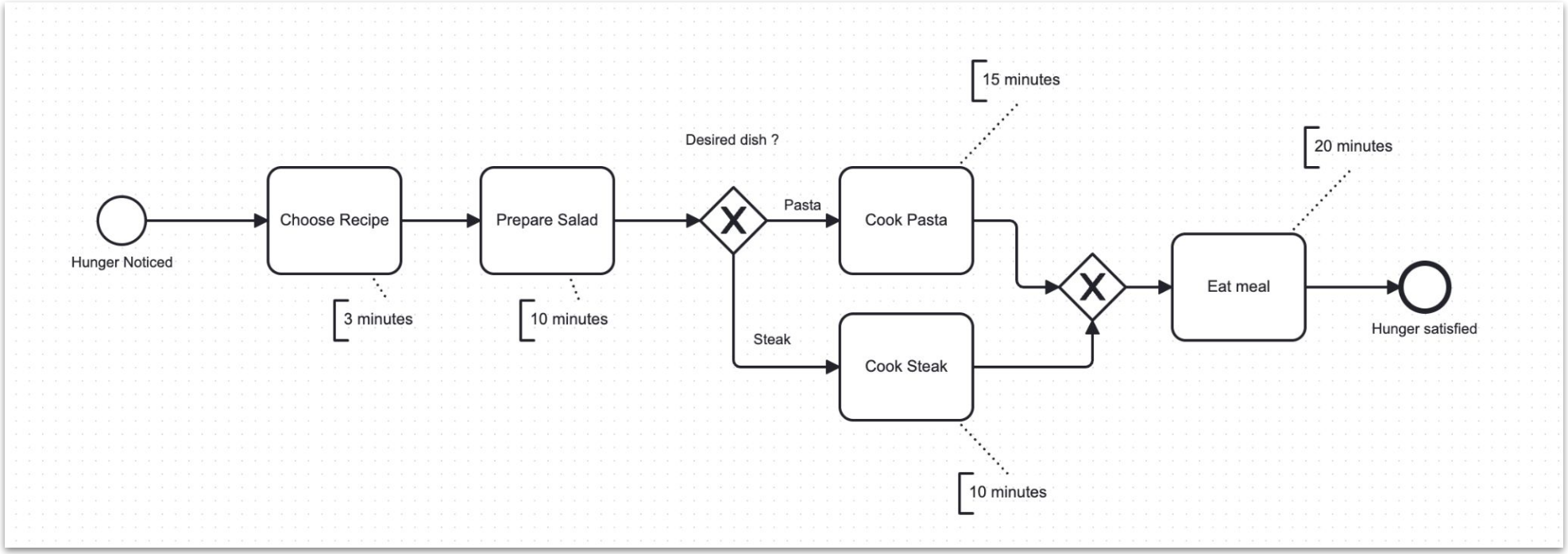
Result

```
PASS ./MapBPMNtoNeo4j.test.js
  MapBPMNtoNeo4j Function
    ✓ should properly convert BPMN data to Neo4j data (17 ms)
    ✓ should handle empty BPMN data
    ✓ should handle invalid BPMN data (1 ms)
    ✓ should execute within a reasonable time frame with large input (2 ms)

Test Suites: 1 passed, 1 total
Tests:       4 passed, 4 total
Snapshots:   0 total
Time:        0.278 s, estimated 1 s
Ran all test suites matching /MapBPMNtoNeo4j.test.js/i.
```

```
PASS ./MapBPMNtoNeo4j.test.js
  MapBPMNtoNeo4j Function
    ✓ should properly convert BPMN data to Neo4j data (17 ms)
    ✓ should handle empty BPMN data (1 ms)
    ✓ should handle invalid BPMN data
    ✓ should execute within a reasonable time frame with large input (5 ms)

Test Suites: 1 passed, 1 total
Tests:       4 passed, 4 total
Snapshots:   0 total
Time:        0.39 s, estimated 1 s
Ran all test suites matching /MapBPMNtoNeo4j.test.js/i.
```



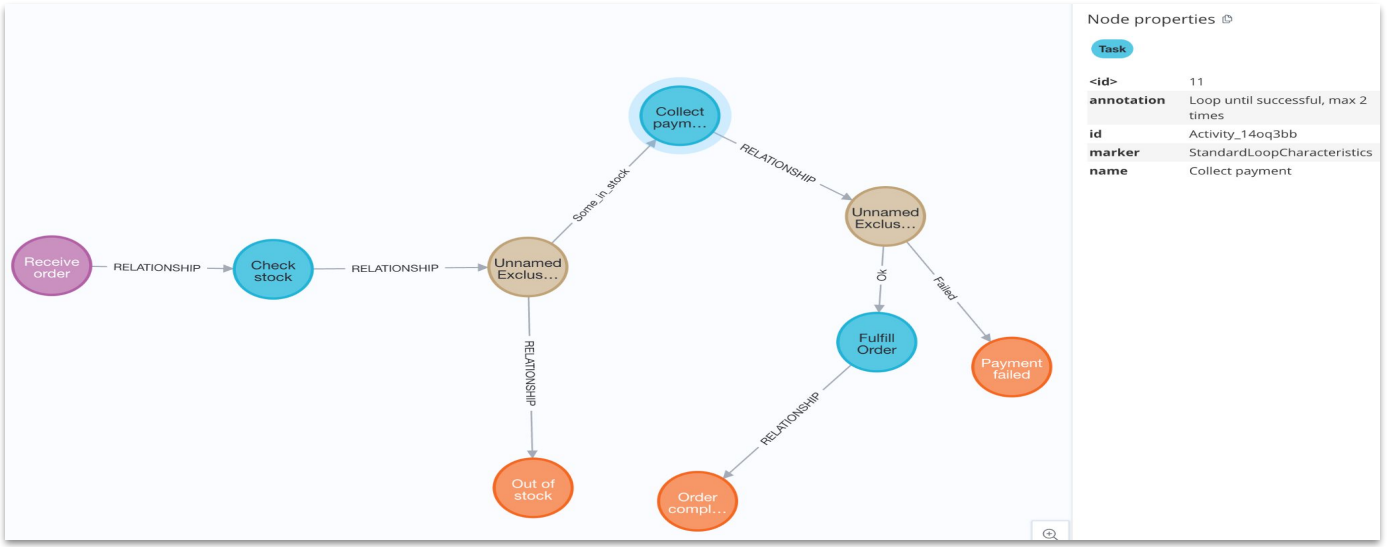
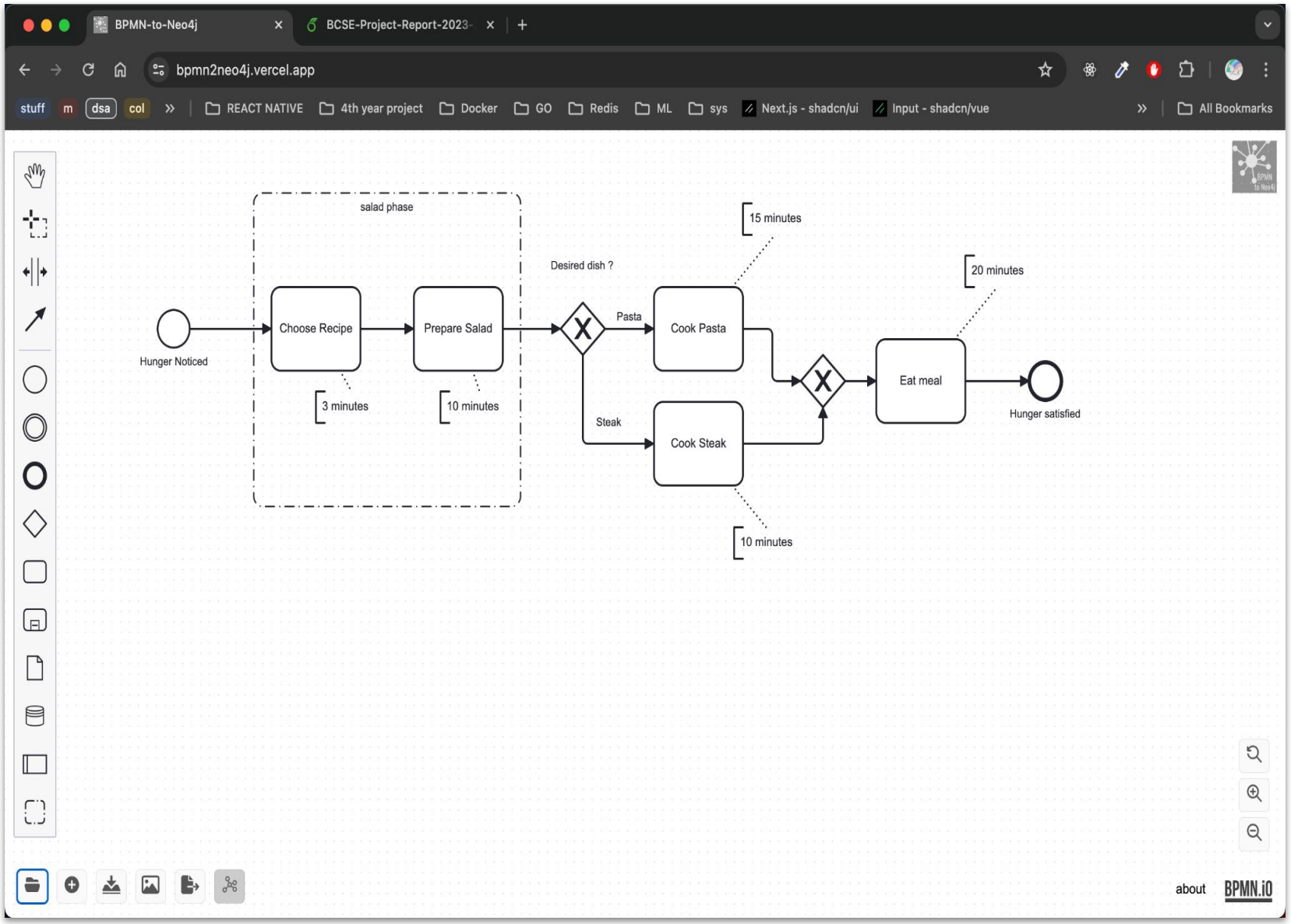


Figure: BPMN Diagram for cooking a recipe.



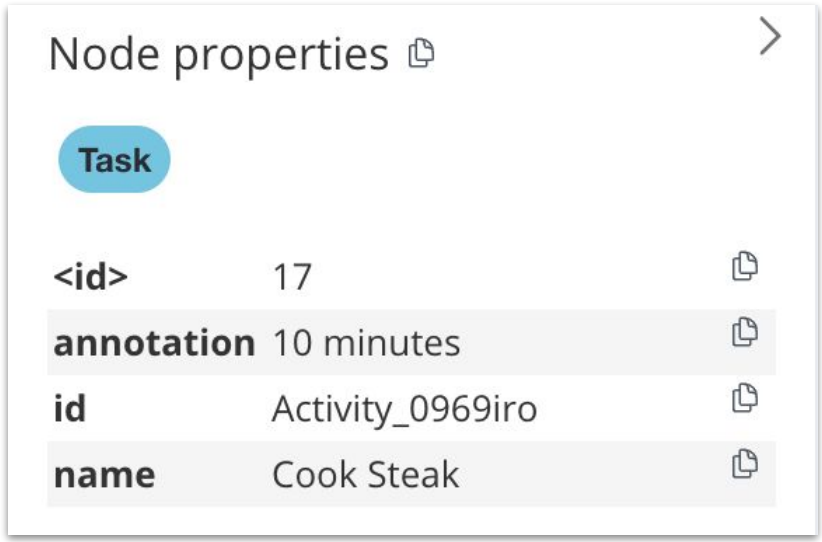


Figure: Showing the cooking diagram in our own BPMN editor and Showing node properties for one node

```
match (n) where n.name=”Cook Pasta” return n;
```

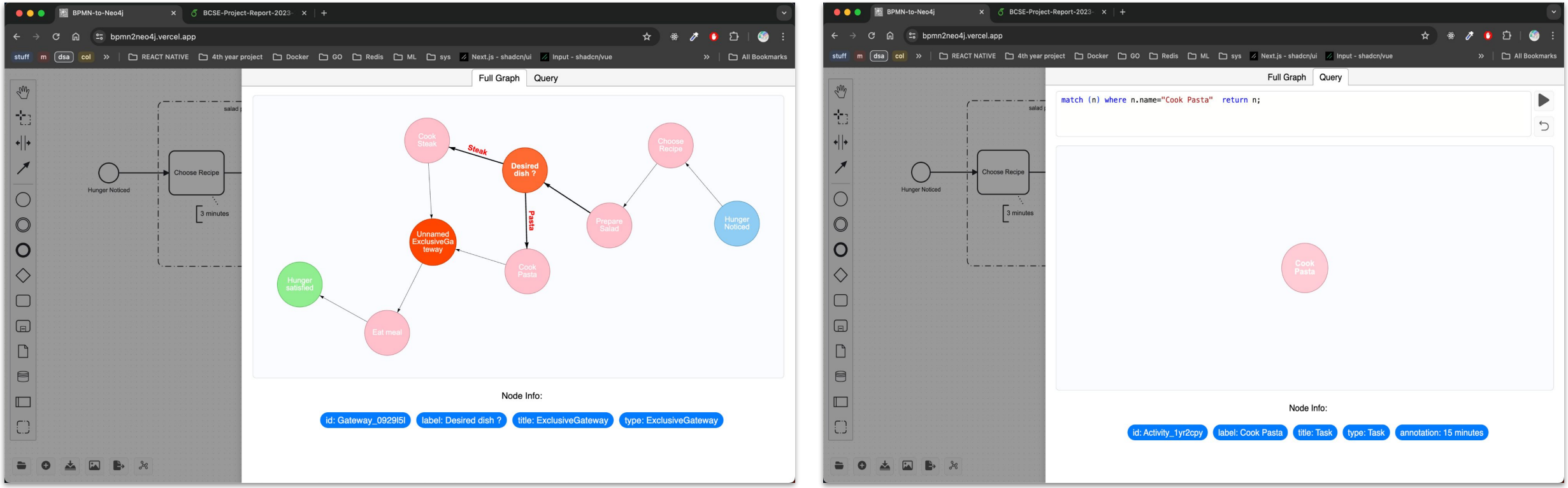
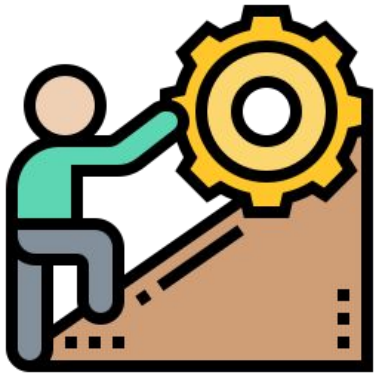


Figure: Showing the graph for the BPMN diagram and Showing the graph after doing a cypher query

Discussion

- This project focuses on converting BPMN 2.0 diagrams into graph structures for database querying and analysis.
- Process involves extracting BPMN data, encoding it in XML, and mapping to a Neo4j graph model.
- Each BPMN element translates into a node, connections between elements are represented as relationships.
- So with the help of Neo4j graph database queries can be done on BPMN diagrams.

Conclusion and Future Work



This project successfully unified the intuitive process modeling of BPMN.js with the powerful data representation of Neo4j, creating a versatile solution for business process management.

By integrating BPMN.js's modeling with Neo4j's robust representation, the project contributes to the evolution of tools and methodologies for more agile, insightful, and interconnected process management in today's dynamic landscape.

Based on this work, we envisage the following areas of research as future work:

- Business Process Optimization
- Automated Knowledge Graph Construction
- Correctness Checking of BPMN Collaborations

References

[1] G. Polančič and B. Orban, “A bpmn-based language for modeling corporate communications,” Computer Standards Interfaces, vol. 65, pp. 45–60, 2019.

[2] P. Pullonen, J. Tom, R. Matulevičius, and A. Toots, “Privacy-enhanced bpmn: enabling data privacy analysis in business processes models,” Softw. Syst. Model., vol. 18, p. 3235–3264, dec 2019.

[3] S. Sakr and A. Awad, “A framework for querying graph-based business process models,” in Proceedings of the 19th International Conference on World Wide Web, WWW ’10, (New York, NY, USA), p. 1297–1300, Association for Computing Machinery, 2010.

[4] F. Duran, C. Rocha, and G. Salaün, “Resource provisioning strategies for bpmn processes: Specification and analysis using maude,” Journal of Logical and Algebraic Methods in Programming, vol. 123, p. 100711, 2021.

[5] S. Bachhofner, E. Kiesling, K. Revoredo, P. Waibel, and A. Polleres, “Automated process knowledge graph construction from bpmn models,” (Berlin, Heidelberg), p. 32–47, SpringerVerlag, 2022.

[6] F. Corradini, A. Morichetta, A. Polini, B. Re, L. Rossi, and F. Tiezzi, “Correctness checking for bpmn collaborations with sub-processes,” Journal of Systems and Software, vol. 166, p. 110594, 2020.

[7] M. von Rosing, S. White, F. Cummins, and H. de Man, “Business process model and notation—bpmn,” in The Complete Business Process Handbook (M. von Rosing, A.-W. Scheer, and H. von Scheel, eds.), pp. 433–457, Boston: Morgan Kaufmann, 2015.

[8] “E2E Technologies Ltd., & 2013. bpmn 0.2.2.” https://www.npmjs.com/package/bpmn.

[9] V. Goulart, V. Serra, M. Stamm, B. Mendes, bpmn.io Admin, N. Rehwaldt, pinussilvestrus, M. Barelkowski, P. Fromme, and M. Trumpf, “bpmn-moddle 8.0.1.” https://www.npmjs.com/package/bpmn-moddle, 2022.

[10] J. Hansson and neo4j organization, “neo4j-driver 5.12.0.” https://www.npmjs.com/package/neo4j-driver, 2023.

[11] M. Kubica, “xml2js 0.6.2.” https://www.npmjs.com/package/xml2js, 2023.