Enterprise Architecture Modeling based on Data Extraction from Business Process Models

Dmytro Orlovskyi
0000-0002-8261-2988
orlovskyi.dm@gmail.com

Andrii Kopp
0000-0002-3189-5623
kopp93@gmail.com

National Technical University “KhPI”
Department of Software Engineering and
Management Information Technology

Problem Description

We present a problem of **Enterprise Architecture** (EA) artifacts extraction from collections of business process models, which organization of higher maturity levels tend to manage.

Nowadays **EA** has at least two definitions upon the context:

1. A formal description of a system or detailed plan of the system at component level to guide its implementation.

2. A structure of components, their interrelations, principles and guidelines governing their design and evolution over time.

EA is expressed using **models** defined by metamodels of various **frameworks**.

Problem Relevance

Modern EA frameworks define the **Architectural Landscape** (AL) as the representation of EA assets that are planned or those are already in use by the enterprise.

Therefore, gathering information about all the most valuable EA assets and their preparing might be a **long-term** and **expensive project**. Moreover, naturally, AL content is constantly evolving as architectural transformation take place. Hence, keeping AL relevant becomes a continuous process.

However, many organizations maintain **repositories** of business process models that serve as a knowledge base.

Research Purpose

The research purpose is to ***shorten time***, ***save costs and efforts*** for gathering information in order to design the **business architecture** view of the AL, which later might be complemented with **data**, **applications**, and **technical architecture** artifacts that support business process execution.

Research Objectives

1. Study state-of-the-art of the problem.

2. Select an EA modeling language.

3. Select a business process modeling notation.

4. Define a mapping between business process and EA modeling elements.

5. Develop the corresponding software.

Formal Problem Statement

Based on the state-of-the-art analysis, we chose de-facto standards of EA and business process modeling: **ArchiMate** and **BPMN**. ArchiMate EA models can be formally described as tuples:

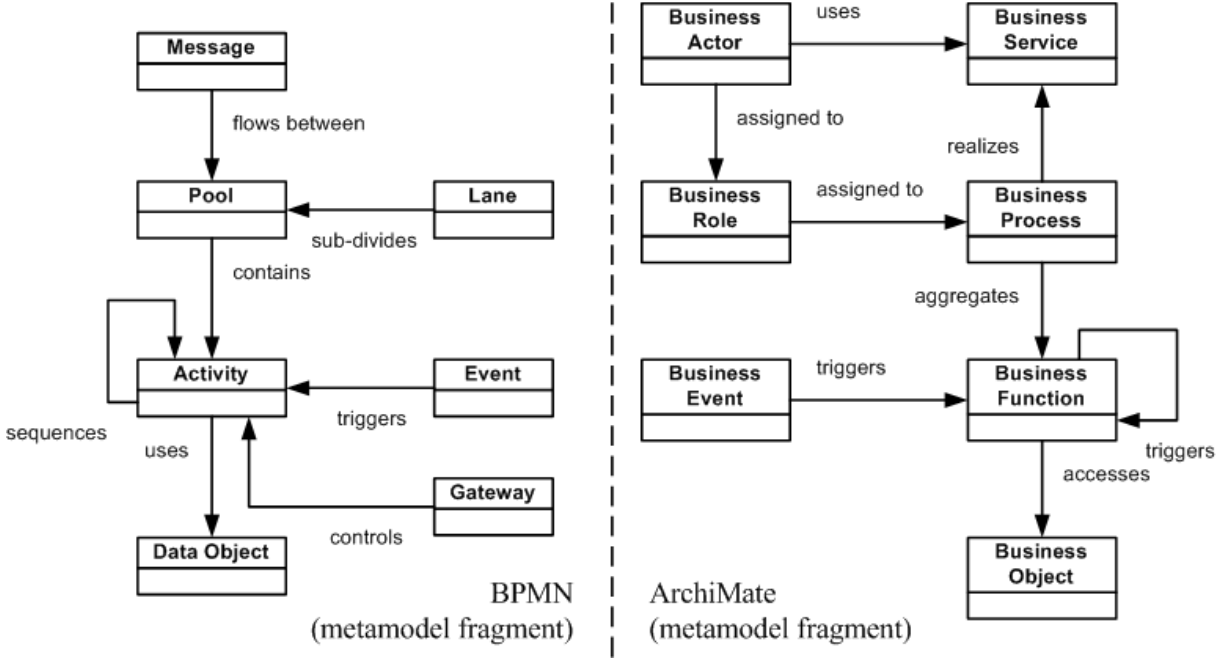
$$AM = <V, E, C, R, vt, et>,$$

V – set of vertices, E – set of edges, C – set of element types, R – set of relations, vt and et – mappings of types to vertices and edges.

Then, subset of V that describes business architecture artifacts should be extracted from a **collection of BPMN models**. As the result we will obtain a ready-to-use ArchiMate EA model with **pre-defined business layer elements**.

Proposed Approach

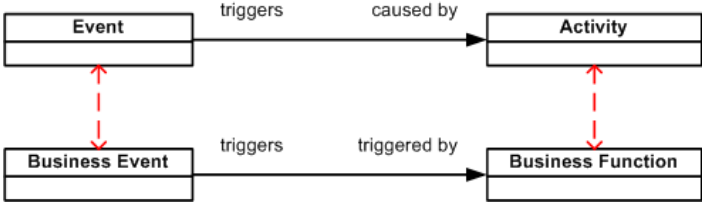
- Proposed approach is based on a mapping between BPMN and ArchiMate **metamodel elements**
- For each of these meta-models we have built **RDF-graphs** (collections "subject-predicate-object" triples)



- Such graphs were queried using **SPARQL** language in order to find similarities and conclude the correspondence between BPMN and ArchiMate models
- **Apache Jena** is a Java-based RDF-storage and API that was used

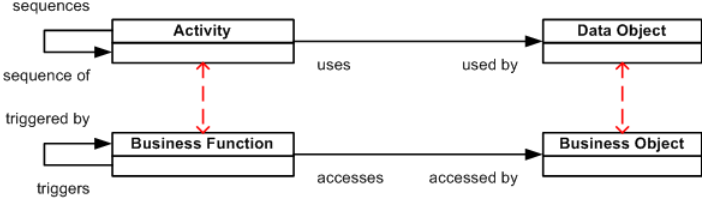
I. BPMN Activity ⇔ ArchiMate Business Function

```
SELECT * WHERE { <Event> ?a ?x . }
SELECT * WHERE { <Business Event>
?b ?y . }
```



II. BPMN Data Object ⇔ ArchiMate Business Object

```
SELECT * WHERE { <Activity> ?a ?x . }
SELECT * WHERE { <Business Function>
?b ?y . }
```

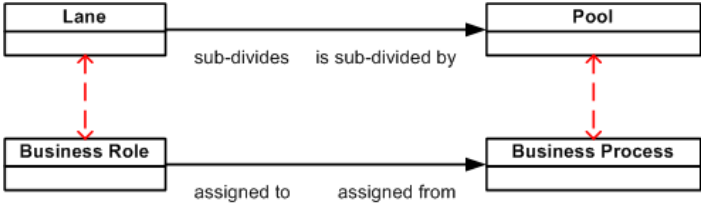


III. BPMN Pool ⇔ ArchiMate Business Process

```
SELECT * WHERE { ?x ?a <Pool> . }
SELECT * WHERE { ?y ?b
<Business Process> . }
```

IV. BPMN Lane ⇔ ArchiMate Business Role

```
SELECT * WHERE { ?x ?a <Pool> . }
SELECT * WHERE { ?y ?b <Business
Process> . }
```



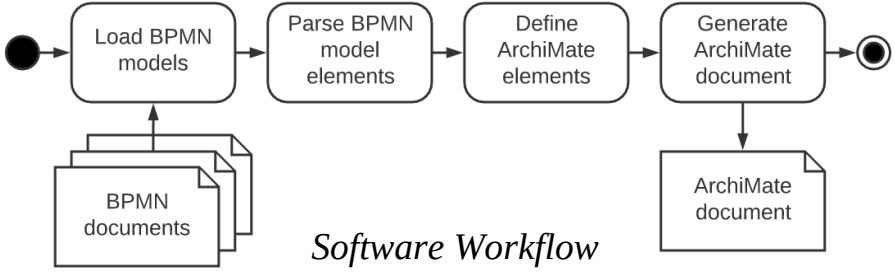
```
RDF-BFS(start_node):
  visited = [start_node]
  queue = [start_node]
  queried = []                    # already queried triples
  while !queue.empty():
    node = queue.poll()
    objects = execute(SELECT * WHERE { node ?x ?y . })
    while statement = objects.next():
      if !queried.contains(statement):
        queried.add(statement) # mark outgoing relation
        if !visited.contains(?y):
          queue.add(?y)
          visited.add(?y)
    subjects = execute(SELECT * WHERE { ?x ?y node . })
    while statement = subjects.next():
      if !queried.contains(statement):
        queried.add(statement) # mark incoming relation
        if !visited.contains(?x):
          queue.add(?x)
          visited.add(?x)
```

In order to formalize such querying of ArchiMate and BPMN meta-models, we used **Breadth-First-Search** (BFS) algorithm that was **modified** for traversing RDF-graphs.

Results of traversing BPMN and ArchiMate meta-models using the **RDF-BFS** method correspond to the results of manual querying.

BPMN	**ArchiMate**
Event *triggers* Activity	Event *triggers* Function
Activity *uses* Data Object	Function *accesses* Object
Activity *sequences* Activity	Function *triggers* Function
Pool *contains* Activity	Process *aggregates* Function
Lane *sub-divides* Pool	Role *assigned to* Process

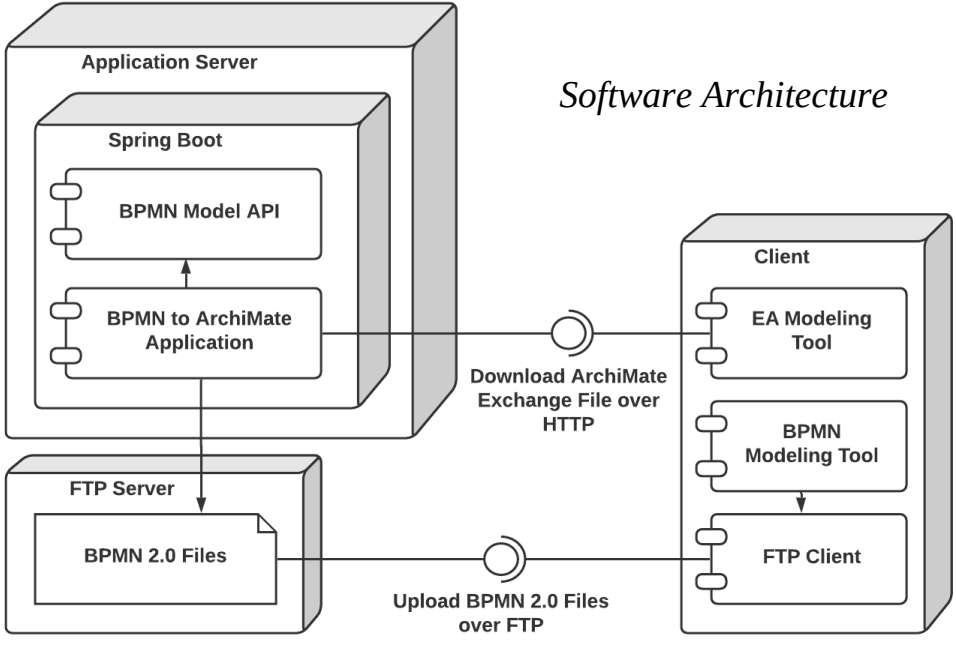
Experiments and Results



Software Workflow

Developed software solution is based on the defined **mapping rules** between BPMN and ArchiMate modeling elements.

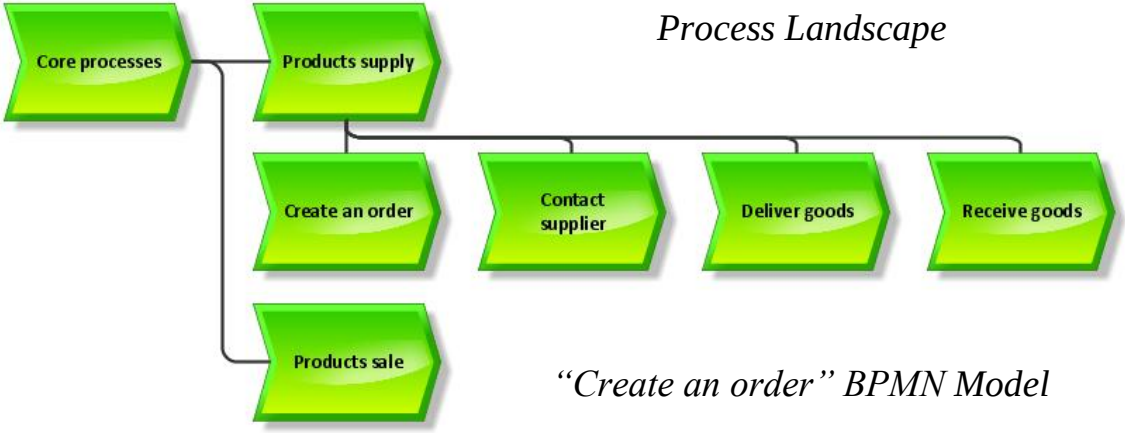
- BPMN models are stored in XML-based documents of **BPMN 2.0** format
- BPMN 2.0 offers **API** to easily extract business process structural elements
- ArchiMate also provides XML-based **exchange file format**



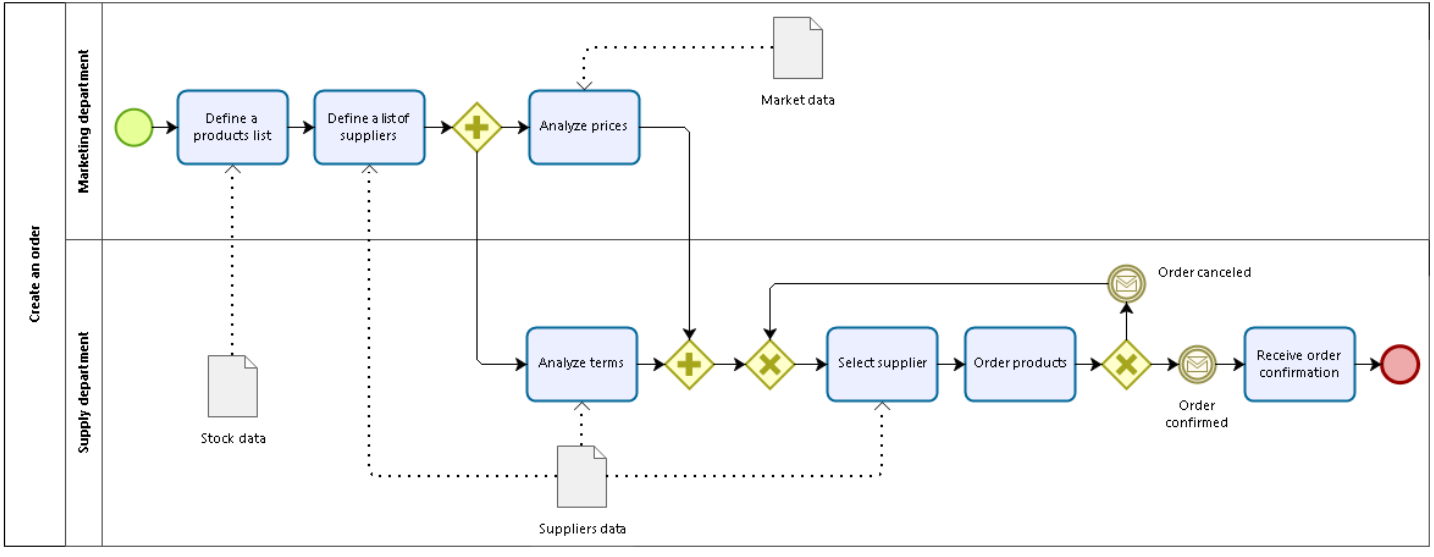
Software Architecture

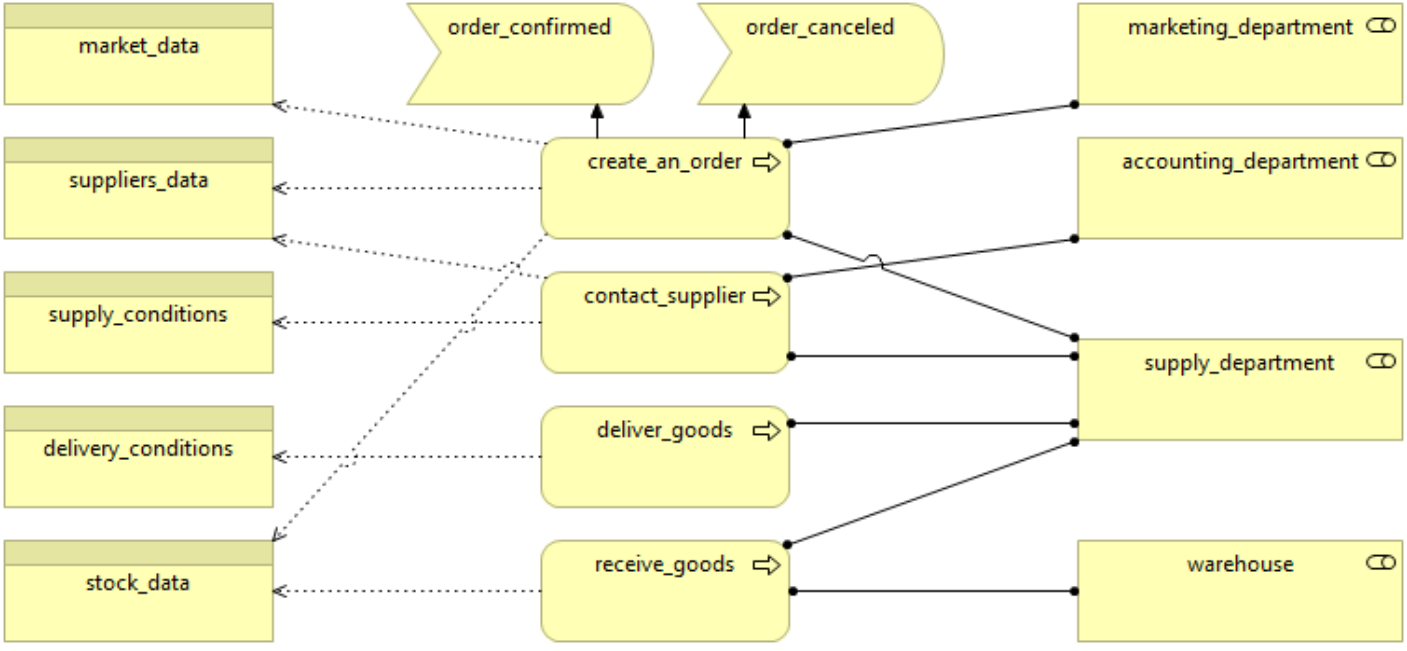
Generic process landscape of **products procurement domain** was used to validate the proposed solution.

Process Landscape



"Create an order" BPMN Model





- Demonstrated business architecture landscape was generated from the set of business process models of the products procurement domain
- Business functions are not included in order to achieve clearer diagram
- Propagation Cost was used to evaluate the AL

- Calculated **Propagation Cost** of 0.10 means that only 10% of the business architecture might be affected in case of required transformation

$$PC = \frac{1}{|V|^2} \cdot \sum_{v \in V} \deg^+(v) = \frac{1}{|V|^2} \cdot \sum_{v \in V} \deg^-(v)$$

Conclusion and Future Work

1. A problem of EA artifacts extraction from business process model collections in order to build AL is described.

2. The approach is based on ArchiMate and BPMN meta-models used to define one-to-one mapping between BPMN and ArchiMate business layer elements.

3. This approach was formalized using the BFS algorithm extended in order to traverse RDF-graphs that were used to store and process BPMN and ArchiMate meta-models.

4. Software solution that was developed to implement proposed approach is web-based and interoperable, since it supports BPMN 2.0 and ArchiMate exchange file formats.

5. Sample set of BPMN 2.0 models that describe products procurement business processes was used in order to validate developed software.

6. Business architecture model extracted from the set of BPMN 2.0 models was evaluated with the help of Propagation Cost metric.

7. Future work includes research in the automatic EA modeling field, including “EA-mining” and related studies.

Thank You for Attention!