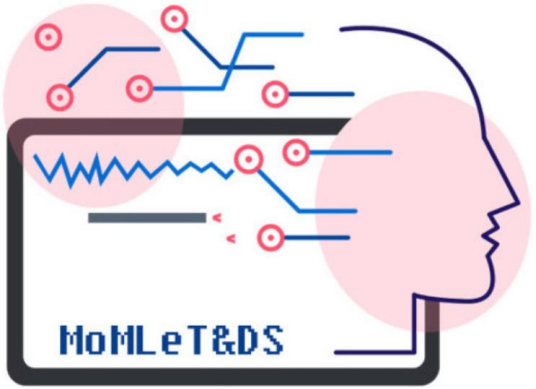
National Technical University
“Kharkiv Polytechnic Institute”

Software Engineering
and Management
Intelligent Technologies



MoMLeT&DS
2022

International Workshop on
Modern Machine Learning Technology
& Data Science

Towards Understandability Evaluation of Business Process Models using Activity Textual Analysis

Andrii Kopp, Dmytro Orlovskyi and Sergey Orekhov
kopp93@gmail.com orlovskyi.dm@gmail.com sergey.v.orekhov@gmail.com

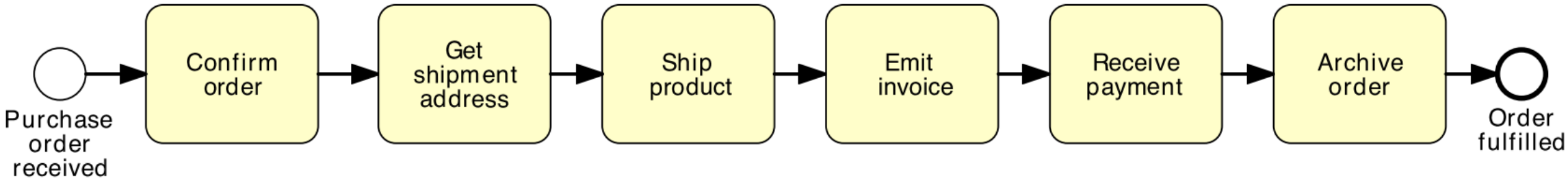
National Technical University “Kharkiv Polytechnic Institute”, Kyrpychova str. 2, Kharkiv, 61002, Ukraine

BPMN Process Modeling Notation

- Currently Business Process Modeling and Notation (BPMN) is a leader and the de-facto standard for business process modeling.
- BPMN models describe workflows as sequences of tasks and events connected using control flows, including start events and end events to signalize beginning and finishing of business processes.

Hence, the simplest BPMN business process consists of events and activities [4]:

- things that happen in an instant are represented by events;
- activities are work units that have a set duration.

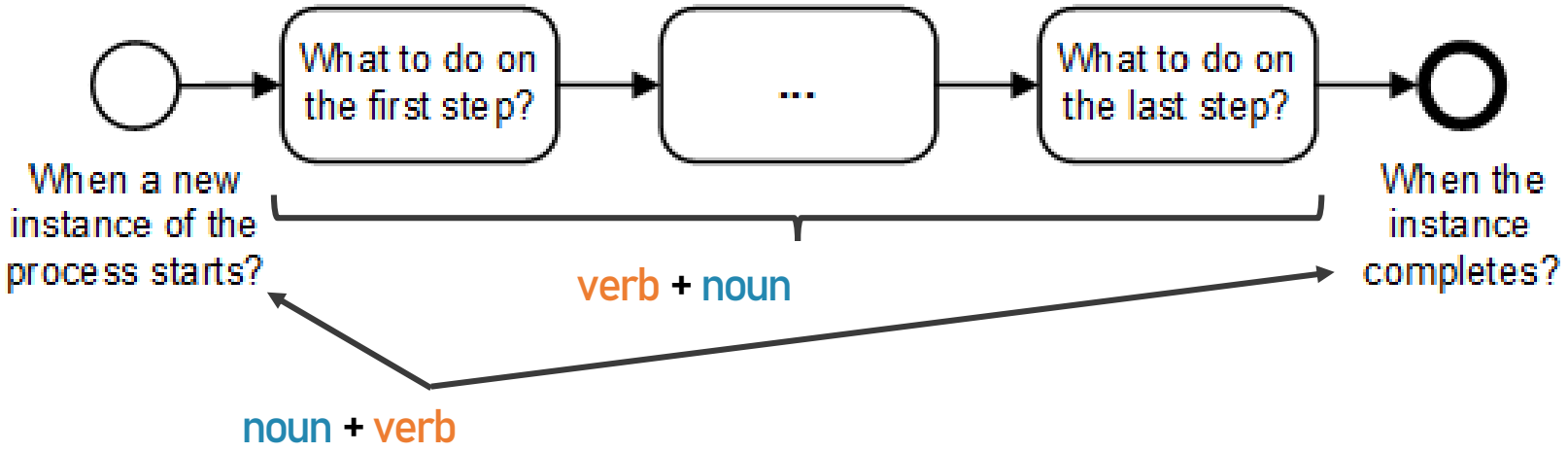


[4] M. Dumas, M. La Rosa, J. Mendling, H. A. Reijers, Fundamentals of business process management, Springer, Heidelberg, 2013. doi:10.1007/978-3-642-33143-5

Business Process Labeling Rules

When describing a business process using BPMN graphical notation, the modeler should answer the following questions:

- "when a new instance of the business process starts?" – for the start event;
- "when the instance completes?" – for the end event;
- "what to do on the particular process step?" – for activities.



Events are usually named as combinations of nouns followed by verbs in past participle form (i.e. "order received", "order fulfilled")

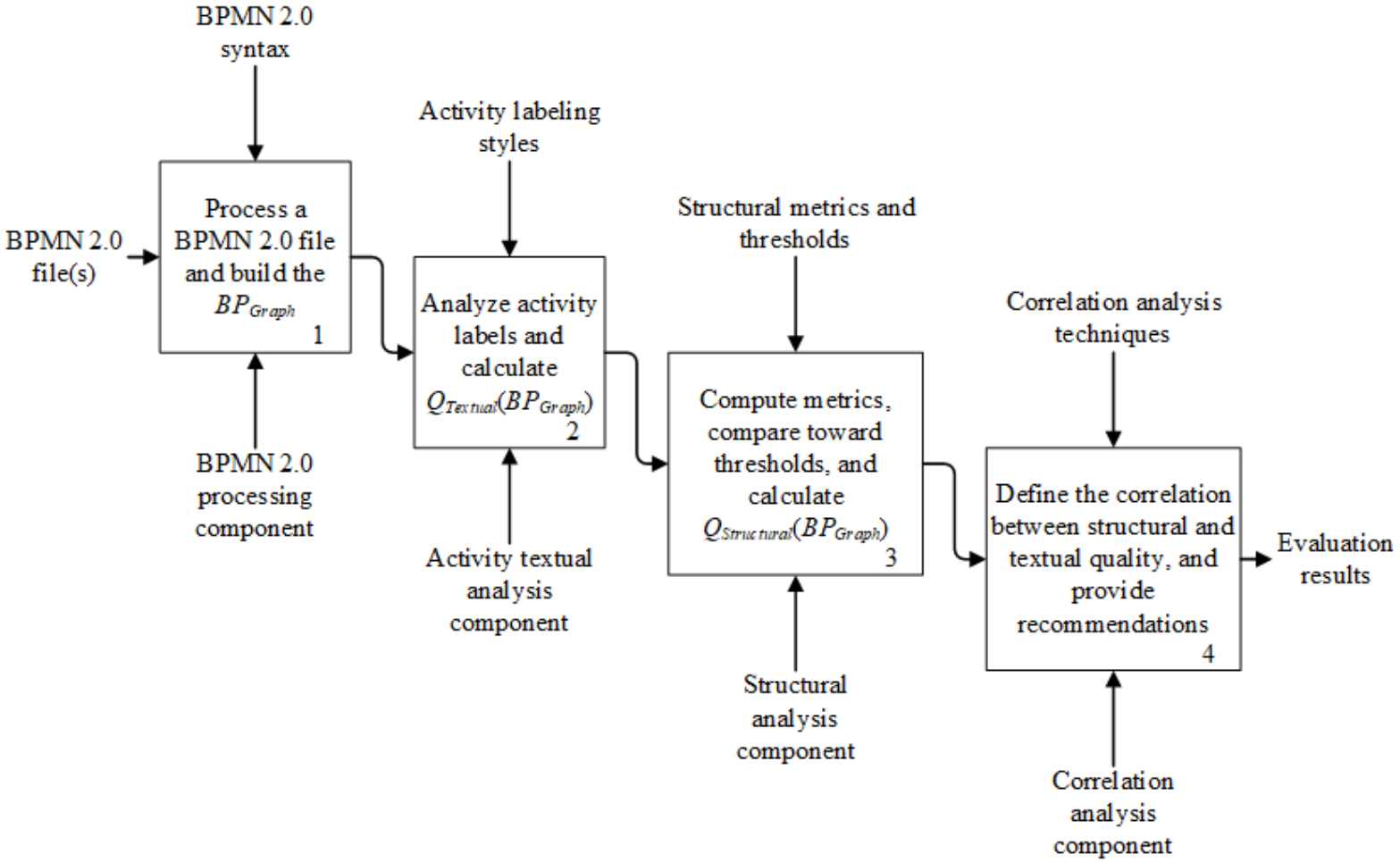
The verb-object labeling style (i.e. a verb in infinitive form followed by the noun: "submit order", "confirm order", etc.) is recommended for activity labels [5]

[5] J. Mendling, Managing structural and textual quality of business process models, International Symposium on Data-Driven Process Discovery and Analysis, Springer, Berlin, Heidelberg, 2012, pp. 100–111. doi:10.1007/978-3-642-40919-6_6

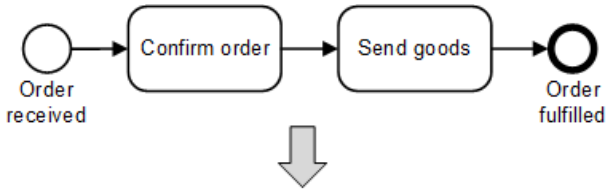
Problem Statement

Poorly designed business process models are hard for understanding and maintenance, and they cannot be efficiently used to document business operations, measure business performance, or find workflow errors that may reduce organizational performance.

Existing studies mostly focus on structural analysis of BPMN model flow using the size and control-flow metrics, and thresholds, while relatively smaller attention is paid to the textual analysis of activity labels used in business process models.



Activity Labels Extraction from BPMN Models

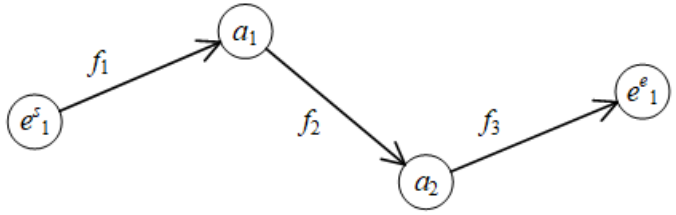



```
<bpmn:process id="Process_0gge2c1" isExecutable="false">
  <bpmn:startEvent id="StartEvent_0qwsd25" name="Order received">
    <bpmn:outgoing>Flow_0bl2fsy</bpmn:outgoing>
  </bpmn:startEvent>
  <bpmn:task id="Activity_002ngbu" name="Confirm order">
    <bpmn:incoming>Flow_0bl2fsy</bpmn:incoming>
    <bpmn:outgoing>Flow_09orffj</bpmn:outgoing>
  </bpmn:task>
  <bpmn:sequenceFlow id="Flow_0bl2fsy" sourceRef="StartEvent_0qwsd25" targetRef="Activity_002ngbu" />
  <bpmn:task id="Activity_1yu2ggg" name="Send goods">
    <bpmn:incoming>Flow_09orffj</bpmn:incoming>
    <bpmn:outgoing>Flow_1y5ye8l</bpmn:outgoing>
  </bpmn:task>
  <bpmn:sequenceFlow id="Flow_09orffj" sourceRef="Activity_002ngbu" targetRef="Activity_1yu2ggg" />
  <bpmn:endEvent id="Event_1g4nvnq" name="Order fulfilled">
    <bpmn:incoming>Flow_1y5ye8l</bpmn:incoming>
  </bpmn:endEvent>
  <bpmn:sequenceFlow id="Flow_1y5ye8l" sourceRef="Activity_1yu2ggg" targetRef="Event_1g4nvnq" />
</bpmn:process>
```

The **<process/>** tag includes all core business process items such as:

- events (i.e. **<startEvent/>** and **<endEvent/>**);
- activities (i.e. **<task/>**);
- and sequence flows (i.e. **<sequenceFlow/>**).

Thus, it is quite easy to read such an XML document and represent it formally using the coherent directed labeled graph:

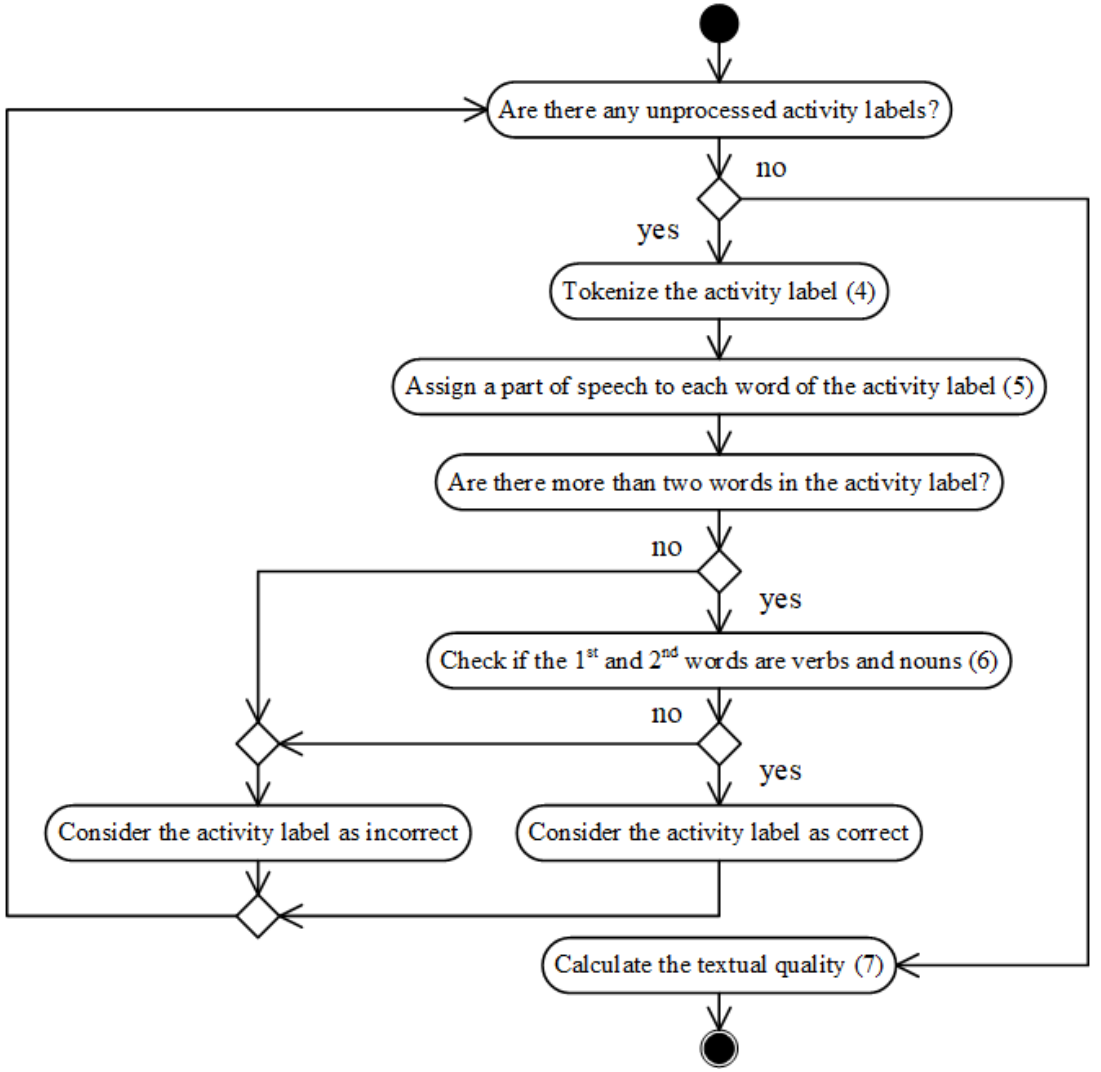


$$BP_{Graph}$$

Activity Labels Analysis Method based on Natural Language Processing

1. Tokenize each activity label to get bags of words that correspond to each of the business process activities.
2. For each word of tokenized activity labels define one or several parts of speech to which it belongs.
3. For each activity label check its length and if the label consists of at least two words, check if the first and second words are verbs and nouns correspondingly.
4. Calculate the textual quality as the ratio between the number of activities, which labels match the verb-object labeling style, and the total number of business process activities.

$$Q_{Textual}(BP_{Graph}) = \frac{1}{Number\ of\ activities} \sum Verb\text{-}object\ activities$$



Structural Analysis of Business Process Models based on Metrics and Thresholds

1. Calculate values of the basic structural metrics to manage the business process model's structural quality:
   - number of nodes;
   - number of start events;
   - number of end events;
   - number of OR gateways.
2. Therefore, using business process modeling guidelines [6], the following threshold values can be defined for the respective structural metrics:
   - do not use more than 31 nodes;
   - do not use more than 2 start and end events;
   - do not use OR gateways.
3. Then, using values of the basic structural metrics and corresponding threshold values, calculate the structural quality as the average of inverse sigmoid function results.

$$Q_{Structural}(BP_{Graph}) = \frac{1}{Number\ of\ metrics} \sum value(metric, threshold)$$

$$value(metric, threshold) = \begin{cases} 1, metric \leq threshold, \\ \dfrac{1}{1 + e^{(metric - threshold - 1)}}, metric > threshold \end{cases}$$

[6] J. Mendling, H. A. Reijers, W. M. van der Aalst, Seven process modeling guidelines (7PMG), Information and software technology 52(2) (2010) 127–136. doi:10.1016/j.infsof.2009.08.004

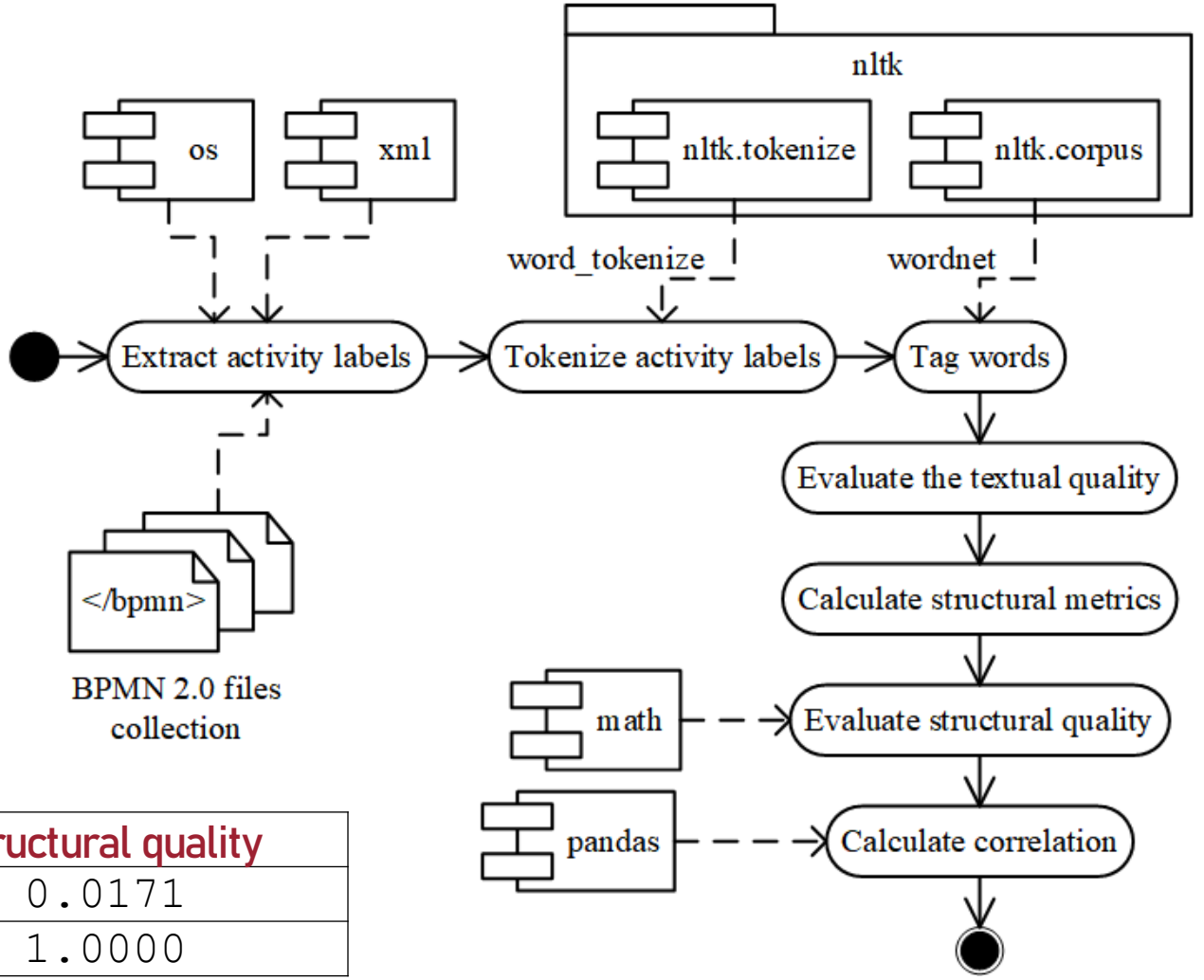
Experimental Results

The dataset taken from Camunda's GitHub repository includes 197 models in English:

- 67 describe the goods dispatch business process;
- 47 describe the insurance recourse business process;
- 34 describe credit-scoring business processes;
- 49 describe self-service restaurant business processes.

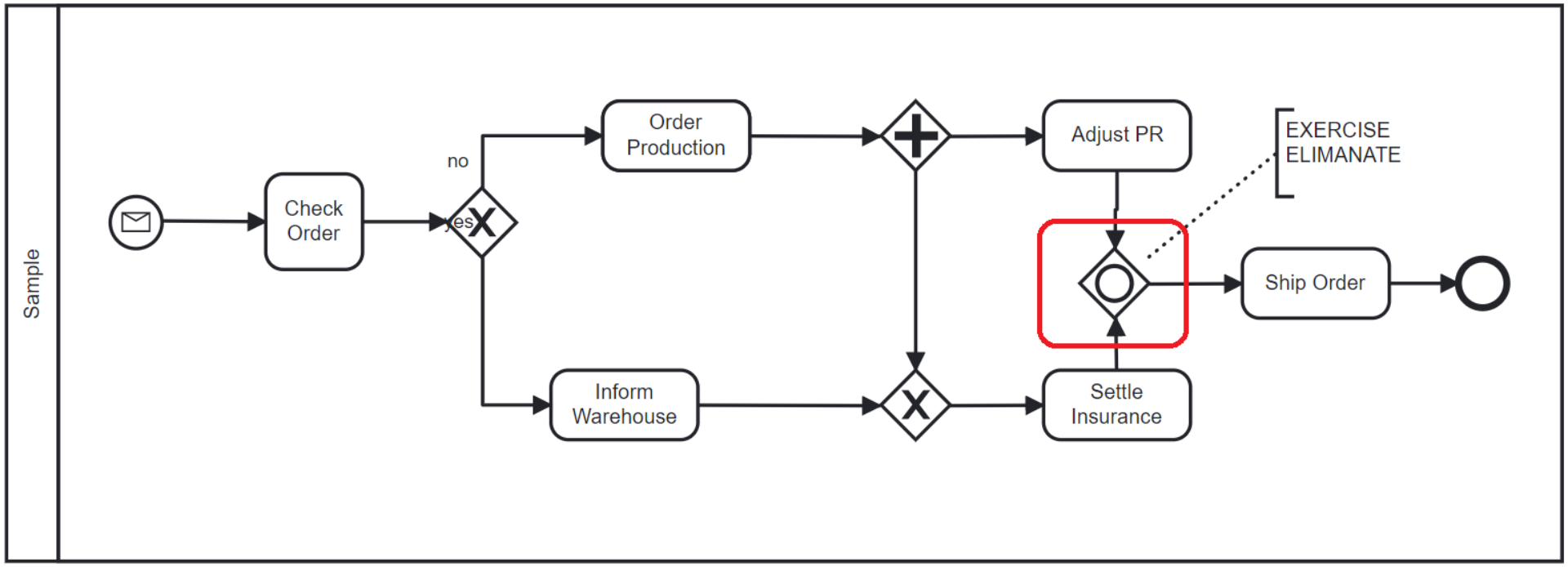
Does the structural quality of business process models affects their textual quality?

➢ There is **no relationship** between textual and structural quality coefficients calculated for each of the experimental BPMN business process models.



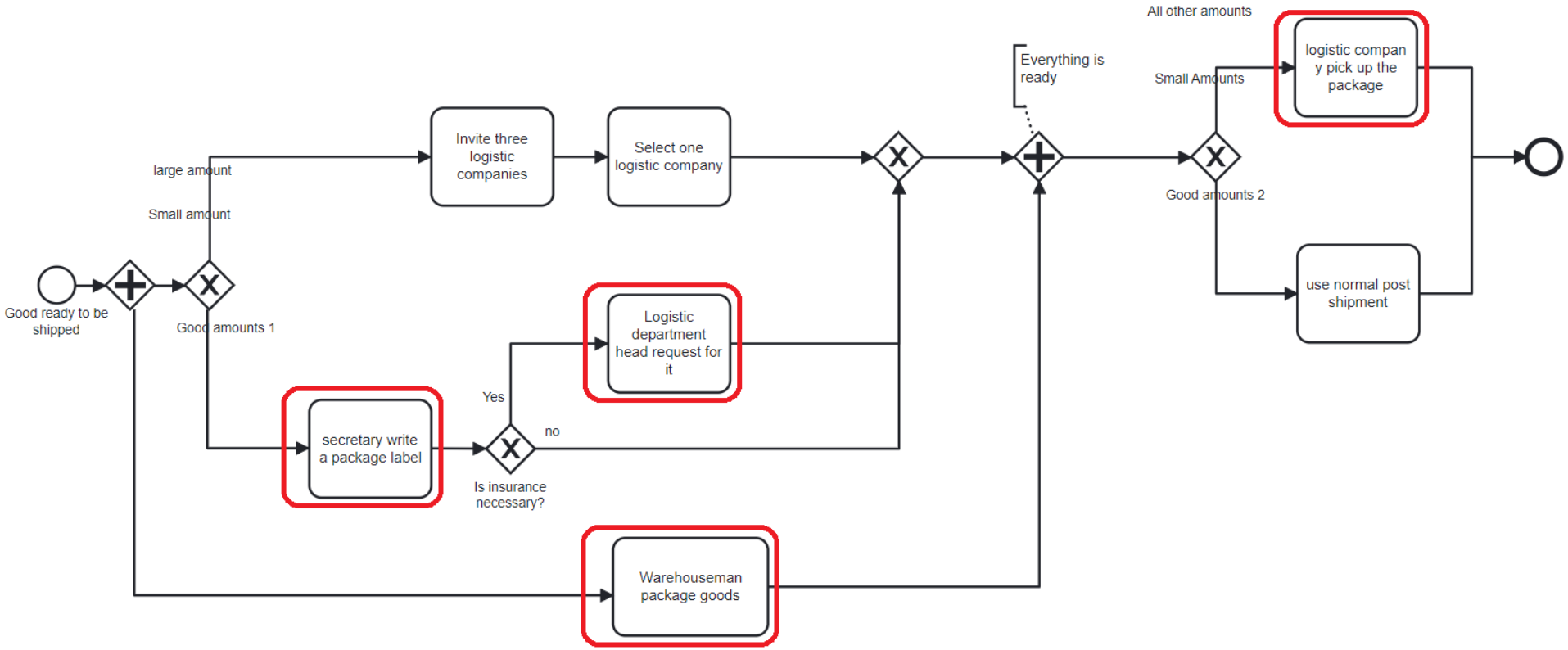
Metrics	Textual quality	Structural quality
Textual quality	1.0000	0.0171
Structural quality	0.0171	1.0000

Example #1



The business process model of high textual quality (1.00) has structural issues (0.88) – the OR gateway is used.

Example #2



The business process model of high structural quality (1.00) has poor textual quality (0.43) – 4 of 7 activities has labelling style that does not match the recommended verb-object style.

Conclusion and Future Work

Findings:

- the structural quality of a business process model does not mean its understandability since there is a bad correlation between these metrics (0.0171);
- demonstrated examples show how the models of high textual quality (1.00) can be of moderate structural quality (0.88) and vice versa – how the models of poor textual quality (0.42) can be of high structural quality (1.00);
- understandable business process models, which are valuable for the stakeholders, should demonstrate high textual and structural quality;
- we recommend business process modelers pay for the textual quality and proper activity labeling as much attention as they pay to the structural quality of business process scenarios;
- having a business process model both structurally and textually sound will make it serve its initial purpose to communicate knowledge about ongoing or planned business processes.

Future work:

- automatic correction of poorly named activity labels;
- continue the study of the relationship between the textual and structural quality of business process models.

THANK YOU FOR YOUR ATTENTION!
QUESTIONS?