Analyzing Model Element Labels of Business Process Model Examples Provided on the Web

Christian Kop
University of Klagenfurt
Klagenfurt, Austria
e-mail: christian.kop@aau.at

Abstract—Business process models are important for Information Systems and the digitalization of the workplace. The Business Process Model and Notation (BPMN) is the de-facto standard in this domain. Therefore, creating BPMN business process models that are understandable to all stakeholders is an important task. The Web provides many business process model examples. These examples are provided by enterprises and consultants who offer technical solutions (i.e., business process modeling tools) or consulting services. Since such models are provided on the Web as introductory learning examples, such examples can also influence novice business process modelers. Therefore, it is worth to examine if such examples have the same quality standards as suggested in the literature. Related to the area of business process modeling, this paper, therefore, focuses on the analysis of such BPMN examples. Particularity, it focuses on the labels of model elements since these labels represent the relationship between the process model and a certain domain.

Keywords—Business Process Model and Notation (BPMN); Business Process Modeling; Quality of Business Process Models; Labels of Model Elements.

I. INTRODUCTION

A good analysis and documentation of business processes is necessary in order to understand the internal behavior of an enterprise and to implement process automation well. The Business Process Model and Notation (BPMN) is the de-facto business process modeling language standard for documenting processes (for the model elements of BPMN, see e.g., the BPMN poster on the Web [16]). It is intended for both high level organizational processes and lower level processes that can be executed by a workflow engine.

For good business process documentation, which is understandable by all stakeholders, skills in business process modeling with BPMN are very relevant. Today, these skills cannot only be obtained by reading books about BPMN or visiting BPMN courses. Instead, it is often much easier and cheaper to click through the Web, looking and reading the diagrams, as well as the enclosed explanations, on a Web site. Thus, Web examples can be taken as surrogates for examples in professional literature (e.g., specialist books). Actually, the BPMN and Process Management (BPM) community (e.g., tool providers and consultants) also have the aim to present BPMN examples on the Web to give either an introduction of the tool features for BPMN modeling or to show modelers how these diagrams look like and how they should be modelled. Hence, a look at such Web examples and their quality for being a standard for novice process modelers can be useful.

There are different aspects of how modeling quality can be defined (e.g., syntactical correctness; adequate drawing of models; adequate color and shape of model elements; adequate labels of model elements, etc.). This work focuses on the labels of model elements. Labels on model elements (e.g., “send application”) as a label example of a BPMN Activity are important since they relate the model to the observed reality. They represent the semantic bearing parts of a domain giving the model elements and thus the whole BPMN model a certain meaning in a specific domain. Therefore, if the labels are not well chosen, a model can be more confusing than understandable and this can lead to a wrong interpretation of models. In the literature, which deals with the quality of model element labels, also recommendations are given how certain model elements should be labelled.

The goals of this work, therefore, are to:
- Check if the introductory learning examples provided on the Web by BPMN experts (e.g., enterprises that offer BPMN tools and consultants offering consulting services) follow the label quality guidelines mentioned in literature.
- Examine if in these examples, the labels are at least well chosen. That means that, even if the labels do not exactly match the guidelines, nevertheless, they make sense in a specific context.

In order to answer this, the analysis of the examples on the Web has been done on a sample extracted from the Web.

The paper is structured as follows. In Section 2, an overview of related work is given and labeling styles together with literature recommendations of good labeling styles are presented. Section 3 describes the preparation of the sample of Web examples for this work. Section 4 focuses on the labels of model elements for business process models provided on the Web. It discusses which kinds of labels are used and compares these with literature suggestions. The paper is summarized in Section 5.

II. RELATED WORK

All aspects of the quality of process models are in the focus of the research community. Issues of deficiencies in BPMN are stated in [14] and [15]. In [3], the author of the book describes how good modeling styles of BPMN should look like. A literature survey about business process modeling quality is given in [5]. Seven guidelines for process
modeling are proposed and verified with user studies in [6].

Some researchers have thought about automating the labeling process of business process modeling and aggregation of process models to support the comprehension of such process models [7] [8]. It was even analyzed how the style, color and arrangement of label parts on a model element improves readability [9] [10].

A. Related Work with respect to Labels of Model Elements

More detailed work on labels of BPMN model elements itself was done in [11] - [13]. These research works are based on data sets of process models from industry. Good labeling styles of Activities, Events and Gateways for three different natural languages were proposed and recommended in [11]. There, violations of these labeling styles are described.

Activities subsume Sub Processes, Event Sub Processes, Transactions, Tasks and Call Activities. In all cases, the tasks of a person or a system, namely, the working step within a process, are described. For the labels of Activities, the following styles were found in this literature:

- **Verb Object Style**: A label that starts with a verb expressing the activity followed by an object, on which this activity is executed (e.g., “create document”).
- **Action Noun Style**: This style has three sub styles: a) A label that has either a nominalized verb only or a compound noun consisting of a verb as the head of this compound noun (e.g., “creation”, “document creation”). b) The Noun can also be a noun phrase with the preposition “of” in between (e.g., “creation of document”). c) Finally, the **Action Noun Style** can also start with a gerund followed by a noun (e.g., “creating document”).
- **The style called “Descriptive” is a style consisting of a subject, a verb in 3rd person singular and an object (e.g., “author writes book”).**

Beside this, there are also labels that do not follow a good style at all. These are labels with nouns only and no verbs at all (e.g., “error”). According to literature [12], the **Verb Object Style** is the most recommended style that should be used for modeling Activities.

With Gateways, a workflow can be divided into several paths, but different paths can also be merged. Most recommended Gateway labeling styles in literature have in common that they should end with a question mark (“?”). Thus, the literature assumes that these kinds of styles are mainly used for XOR and OR Gateways since in these Gateways a decision is made, which can be expressed as a question. On contrary, an AND (parallel) Gateway does not need such a label since no decision is made. Such questions in Gateway labels can be expressed in one of the following styles:

- **Question with Noun and Verb in Past Participle** (e.g., “document created?”)
- **Infinitive verb question** (e.g., “approve contract?”)
- **Object with adjective question**: A phrase consisting of an object followed by an adjective or an auxiliary and an adjective (e.g., “parts available?” or “parts are available?”)
- **Equation question**: A phrase consisting of an object followed by a logical operator and a value (e.g., “amount is greater than $ 200”).

Even here, there are labels, which are treated as bad labels. As a counter example for good quality, a noun only (e.g., “result?”) is treated, since from this kind of label it is not possible to derive a clear decision. For Gateways, the most recommended label style is: **Question with Noun and Verb in Past Participle** [12].

Finally, events that can occur within a process are modelled with the model element Event.

Labeling styles for Events can be classified as followed:

- **Verb in Past Participle Style**: This can be characterized by an object followed by a verb in past participle or followed by a (modal) auxiliary and a verb in past participle (e.g., “document created”, “document has been created”, “document must be created”).
- **Predicative Adjective Style**: Here, a noun together with a predicative adjective is used to label an Event (e.g., “document correct” or “document is correct”).
- **Categorization Style**: Two nouns are related with a verb (mainly the verb “is”) in order to express that the term specified with the first noun can be categorized according to the term expressed with the second noun (e.g., “person is author”).

Modelers also use labels that better should not be used for Events at all, since they do not provide sufficient information to a reader. For instance, they use a noun only (e.g., “inquiry”). The **Verb in Past Participle Style** is the one which is most recommended as a labeling style for Events [12].

Beside simple labels, it has also been examined in literature that modelers use complex phrases and sentences for Activity labels instead of drawing more model elements with simpler, so called canonical labels. Especially in [13], these kinds of inconsistent use of labeling, so called non-canonical patterns, are examined. Three categories of complex non-canonical label patterns were detected:

- **Complex control flow label**: The label of an Activity consists of a sequence of verbs, each describing an Activity, which are concatenated with “or” or “and”. This verb sequence, however, implicitly expresses a decision (in the case of “or”) or a parallel respectively a sequential execution of several Activities (in the case of “and”). It does not express an atomic working step. Thus, instead of one Activity with a complex label, several Activities with simpler labels together with control flows can also be used. Other complex labels of that kind are phrases, which end with “as required”, “as if needed”, as well as sentences or phrases expressing an iteration (e.g., “while … “, “as required… “, “for each …”).
- **Extra specification of data, resources and time**: In this category, the label of the model element not
only contains the necessary information, but also additional information that is often given in some sorts of brackets (e.g., “clear differences (inventory management”) ). Most often, either this extra information should be itself explicitly modelled with a model element like an Event, Activity or Gateway or this extra information is useless.

- **Implicit Action and Decision**: Here, the label and the modelling element do not fit. For instance, the label of an Activity is expressed in terms of a pattern that is typically used for an Event (e.g., “Order received” instead of “Receive Order”).

In literature, these categories of non-canonical labels are seen as patterns that can confuse the reader of a model.

B. **Focus of this Work with respect to Related Work**

In this work, the labels of the model elements are also examined. The focus is on model elements of process diagrams. An explanation of process diagrams can be found at [4]. In addition to previous work in the related literature, the main emphasis of this work can be characterized as follows. Instead of working with data sets from industry, the aim of this paper is to look for BPMN examples on the Web. Existing results of labeling guidelines in literature are taken as a reference. With this as a basis, the Web examples are examined and compared with the given guidelines.

Furthermore, for analyzing the labels, this work does not only consider Activities, Gateways and Events as such, but also explores different types of Activities, Gateways and Events in detail.

III. **Preparation of the Work**

In order to check how different enterprises, which sell BPMN modelling tools, as well as consultants, provide BPMN diagram examples, the following procedure was executed to get the sample. In the first step, the search term “BPMN” was entered into the search field of Google. This search engine was used as a means to choose the sample. In order to get diagrams first and not descriptions of BPMN, the image result list of the search engine was used. Here, it was expected to get various images of BPMN diagram examples. Once the list of diagram images were generated by the search engine, in the second step the list was manually examined. For each image, it was first of all decided if this image is really a BPMN diagram example in English provided at a Web site or if it is not. If it was indeed such a diagram, then the link to the respective Web site, from which the search engine listed the image, was collected. For this purpose, the link was entered into a file in order to generate a list of Uniform Resource Locator (URL) links. At the end of this URL collection step, a list of URL links, each containing at least one image of a BPMN diagram was collected in the file. In the third step, the file with these links was further examined. For doing this, the file with the URL link list was automatically scanned and each link was grouped to a Web domain.

Afterwards in the fourth step, each link, as well as the additional link to the more general Web domain, was once again examined further. From all these sources, images of BPMN diagram examples were downloaded and collected on the local file system. The aim of this fourth step was to find more BPMN images provided at this Web domain. Much more images were found and collected.

In the fifth step, all these collected images were manually examined and the individual BPMN model elements together with their features and labels for each diagram and domain were transcribed into a database. This data set was then analyzed according to the aim of this work.

In total, 43 Web domains and the BPMN 2.0 by Example document of Object Management Group (OMG) [2] were examined. The BPMN 2.0 by Example document was included, since this is also an important information resource about BPMN on the Web. During this collection phase, images, which were not readable, were filtered out. Images that are used as BPMN counter examples (i.e., how a BPMN diagram should not be modeled) were filtered out too, since the focus is on models that are seen as correct by the provider. The remaining examples, which in total are 346 diagram images of BPMN model examples were used for this work. Furthermore, only distinct labels were analyzed. This should avoid that a certain label pattern appears too often just because the same label (e.g., “order product”) is used in many examples.

IV. **Labels of Model Elements**

In this section, the analyzed distinct labels of Activities, Events and Gateways specified in process model examples on the Web are discussed with respect to the recommendations in literature mentioned in Section II. These model elements are chosen since a) quality guidelines as references already existed in literature and b) introductory examples hardly make sense without them. Especially, this holds for Tasks as a subset of Activity and Events. Gateways on the other hand have to be used as soon as a process model does not have only a single sequence, but the specified process in the process model branches to several paths. Thus, in most process models, except the most trivial once, Gateways are important. Furthermore, these labels are analyzed in detail according to the specific model element, since for different model elements different labeling strategies are needed.

A. **Labels of Activities**

Activities represent those parts of a process where somebody or something should act in order to progress the process. Therefore, an active verb, which is the best word category for acting, should be used to label these model elements. In literature, the Verb Object Style is preferred. An object itself could be a noun (simple object compound) or a noun phrase.

In the sample, 944 distinct Task labels were found. The majority of these distinct Task labels (74.6 %), which were detected in the diagrams, follow this Verb Object Style, where the object is a noun and the direct object of the verb (e.g., “specify vacancy”, “ship item”, “review results”). In some cases, an article is added (e.g., “select a pizza”). Only in 2.3 % of all cases, a single verb or a verb together with an adverb is the only label for a Task (e.g., “publish”, “rate
negatively"). 12.9 % of the Task labels extend the suggested Verb Object Style a little bit. In these labels, the object is a noun phrase (e.g., “nomination form” in “send nomination form”). Also, cases were found, where the object is an indirect object (e.g. “communicate to customer”) or there are two objects (direct and indirect object) following the verb (e.g., “deliver books to customer”). In 10.2 % of the Task labels, the model designers used other label styles for Tasks. For instance, they used nominalization of a verb (e.g., “delivery”), they used full sentences (e.g., “why have you bought so many sticks of sausage?”), they concatenated verbs (e.g., “add paperwork and move package to pick area”) or they used a condition phrase (“check if extra insurance is necessary”). It seems that model designers also like to specify many verbs for the label of a Task. To summarize, the Verb Object Style preferred in literature is also used in the majority of cases on the Web (Figure 1).

Figure 1. Percentage of Task label styles.

For Sub Processes, the situation is as follows: From the 85 distinct labels of normal Sub Processes, 42.4 % have a nominalization of the verb (e.g., “ordering”, “creation”) as their label (i.e., Action Noun Style). In 55.3 % of the cases, Sub Processes follow the Verb Object style. The rest either does not have a label or it is a complex expression (e.g., “send out application forms & reminders”). Hence, no definite labeling preference can be found in these examples from the Web.

There are not enough Event Sub Processes and Transactions in the sample. Therefore, here it is hard to make a good proposition. In these few examples, the labels follow the Verb Object Style. There are also not so many Call Activities in order to make a proposition. It could only be observed here that the model designers of the Web examples much more often used a nominalization of a verb rather than the Verb Object Style.

B. Relationships between Labels and Specific Task Types

Since about a fifth of all modeled Tasks are modeled as User Tasks, it is interesting what is modeled as a User Task. Especially, it is interesting what is modeled as a User Task in comparison to what is modeled as a Manual Task. Therefore, the labels of the two Tasks are analyzed.

From the point of view of the BPMN specification [1], there is a clear distinction between a User Task and a Manual Task. A User Task is performed by a human, but assisted by a business process runtime. A Manual Task is also performed by a human, but without assistance of a business process runtime.

It could be expected that a task that is supported by a system uses other verbs in the label, than a task that is purely done by a human. However, according to the labels it is not always possible to differentiate between a User Task and a Manual Task. Of course, labels with a verb were found that fit with the purpose of a User Task (e.g., “edit 1st level ticket”, “fill in purchase form”, “book flight”, “find student’s position”). On the other hand also labels were found, which do not perfectly fit with the purpose of a User Task (e.g., “hire staff”, “plan interview”, “read book”, “rent office”, “ship book”, “train new employee on job specifics”, “discuss nominations”, “announce Nobel prices laureate”, etc.). The labels for User Tasks and Manual Tasks are set arbitrarily. One interpretation could be that it is the modelers decision to see something as a Manual Task (without process engine support) or a User Task (with process engine support) and it depends on the purpose of the model (i.e., whether it is a workflow model or not). Particularly, a User Task can be more than a simple user interaction with the Information System. Thus, if a workflow for a workflow engine is specified with BPMN then it seems that every Manual Task can also become a User Task. A second interpretation could be that modelers of these introductory learning examples do not really want to distinguish between User Tasks and Manual Tasks at all. Therefore, they prefer to model a User Task even in a situation where a Manual Task would be the right choice.

The frequency of other Task types is very low and, except for Send Tasks, no relationship between labels and these Task types were found. For the 38 distinct labels of Send Tasks, in this sample it turned out that 52.6 % of the distinct Send Task labels start with the verb “send”. Further, 26.3 % have a verb like “email”, “inform”, “notify”, “distribute”, “post”, “submit”, “order”. All these other label examples can be seen as variants of sending. Thus, it could be concluded that labels of a Send Task are in accordance with the purpose of this Task type.

C. Labels of Events

When talking about labels on Events, firstly, it has to be examined if Events have labels. While BPMN modelers always give labels for Tasks, they are not so systematic if they have to specify labels for Events. From all the Start Events found in the diagrams of the sample, 45.6 % do not have a label. From these, most of the Events (85.7 %) are untyped Events (i.e., Events that are not further classified to be specific types of Events). However, a few cases were also found with Link Events, Message Events and Signal Events that have no label.

For Intermediate Events, fewer cases with no labels exist. Only in 13.7 % of all Intermediate Events no labels were detected. Particularly, the Intermediate Timer Event and Message Event are those types with no labels. These two Event types also had a high frequency within the Intermediate Events types. 37.6 % of all Boundary Events do not have a label. 55.6 % of all End Events do not have a
label. From these End Events with no label, 80.8 % are untyped.

To summarize this, for the examples provided on the Web, the modelers responsible for these examples especially do not see the necessity to label Start and End Events. Particularly, this happened if these Start and End Events are untyped Events. Unlabeled model elements, however, cannot be understood well. If novice process modelers see such unlabeled model element examples on the Web, they might take it as a standard although it should be avoided.

After the examination of Events with no labels, it is interesting to see what kind of labels Events have. It is expected that specific Event types have specific types of labels. For example, Message Events and Timer Events are labelled in different ways. For this analysis, six Event types were further examined, since these six Event types cover 87.2 % of all Event types in the sample. These Event types are: Timer Event, Message Event, Signal Event, Compensation Event, Terminate End Event and the untyped Event.

All labels of the Timer Event have, of course, in common that they specify time. However, this is done in various ways. Table 1 presents a list of representative Timer Event labels. In this list, the grouping of the individual labels, suggests label patterns of similar structure.

<table>
<thead>
<tr>
<th>TABLE I. TYPICAL LABELS OF TIMER EVENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>• wait until next business day</td>
</tr>
<tr>
<td>• 24h; 10 min; 60 minutes; one week; 2 weeks; 24 hours; 14 days; 48-hours</td>
</tr>
<tr>
<td>• september year n-1</td>
</tr>
<tr>
<td>• wait 6 days; wait some time; wait until Thursday, 9am</td>
</tr>
<tr>
<td>• 1st day of month; 20th of each month</td>
</tr>
<tr>
<td>• 3 business days</td>
</tr>
<tr>
<td>• friday at 6 pm pacific time; friday, 6 pm pacific time</td>
</tr>
<tr>
<td>• timeout; time out (1 week); order timed out</td>
</tr>
<tr>
<td>• content expired (5 days)</td>
</tr>
<tr>
<td>• delay 6 days; delay 6 days from announcement</td>
</tr>
<tr>
<td>• &lt; 60 min; &gt; 60 min</td>
</tr>
<tr>
<td>• expires at set deadline</td>
</tr>
<tr>
<td>• auction over</td>
</tr>
<tr>
<td>• 10 min wait</td>
</tr>
<tr>
<td>• 12 o'clock</td>
</tr>
<tr>
<td>• start time; finish time</td>
</tr>
<tr>
<td>• on next Wednesday</td>
</tr>
<tr>
<td>• start on Friday</td>
</tr>
<tr>
<td>• every 10 minutes; every 24 hours</td>
</tr>
</tbody>
</table>

From the examples, it can be seen that they are not in accordance with the Event labeling style recommended in literature (Verb in Past Participle Style). Nevertheless in the context of a Timer Event, many of these labels make sense.

For a Message Event, it has to be distinguished between a throwing Message Event and a catching Message Event. Usually, it could be expected that a catching Message Event follows the Verb in Past Participle Style. However, the found catching Message Events have a greater variety. Beside the typical Verb in Past Participle Style also catching Message Events were found that consists of

- a noun (compound noun) or noun phrase only - i.e., the message (e.g., “payment”, “complaints to customer service”)
- a complete sentence (e.g., “where is my pizza”)
- a verb in past participle only (e.g., “paid”).

- a noun with an adjective (e.g., “assignment complete”)

Unfortunately, not so much throwing Message Events were found. Most of these throwing Message Events had no label at all. On the basis of the remaining throwing Message Events with labels, it can be said that throwing Message Events use the Verb Object Style as it is usual in Task labels. Since a throwing Message Event can be used as an alternative for a Send Task, this label pattern makes sense, although literature recommends the Verb in Past Participle Style without considering the specific type of Event.

The found Signal Events (catching or throwing) follow the Verb in Past Participle Style to a large extent. Beside this, the following interesting label examples were also found: “on alert”, “undeliverable”. According to literature, these examples would be out of the scope of the recommendations.

Many of the Compensation Events do not have a label. Those that have a label mainly follow the Verb Object Style (e.g., “cancel reservation”, “undo book travel”). Many of these labelled Compensation Events are throwing Events. Once again, this is out of scope of the recommendation in the literature, which in general prefers the Verb in Past Participle Style for Events. But, in this special case of throwing Events, which rather express an active action than a passive reaction, the Verb Object Style makes sense.

Most of the Terminate End Events do not have a label. The few remaining Terminate End Events with labels follow the Verb in Past Participle Style or just have the label “terminate” or “end”, respectively.

The analysis of untyped Events is split into the analysis of Intermediate Events, Start Events and End Events. No untyped Boundary Events with labels appeared in the sample. This analysis provides the following results. Untyped Intermediate Events follow the Verb in Past Participle Style. The labels of untyped Start Events do not only follow this style. Instead, some of them only have

- a noun, compound noun or noun phrase - i.e., the object only (e.g., “application”, "existing process"),
- an adjective (e.g., “hungry”) or phrases starting with an adjective (e.g., “hungry for pizza”),
- a simple sentence (e.g., “the store opens”).

The labels of untyped End Events follow the Verb in Past Participle Style to a large extent.

D. Labels of Gateways

The labels of Exclusive and Inclusive Gateways vary. The style Question with Noun and Verb in Past Participle is not the only one. Again, additional patterns exist: Objects only (i.e., nouns, compound nouns and noun phrases), verb in past participle only, state of an object (i.e., where the state is represented by an adjective or by the word “ok”), comparison with operators (e.g., “>”, “<”) or with words (e.g., “above”). What is common to many labels is the character “?” at the end of the label. Many Exclusive Gateways and Inclusive Gateways even do not have a label although they branch the process path. Such cases once again can be seen as a contradiction to the recommendations in literature.
Beside the label of the Gateway itself, it is also important to analyze the labels on the Sequence Flows, which leave the Gateways. About a fifth of all these labels have either the value “yes” or “no”. The rest varies (e.g., “1”; “>=20”; “40 %”; “2nd level issue”; “50 % education training”; “all items available”; “allow extension”; “bicycle costs >= 500 usd”; “capacity & parts available”; “capacity not available”; “capacity ok”; “employee is ready for work”; “fix in release”; “in stock”; “is junk mail”; “no more responses”; “not accepted”; “payment received == false”; “purchase 1”; “put on hold”; “ready with request”; “simple”). Process Diagrams intended for workflows also have Gateway labels like “${order.price <= 250}” or “${!approved}”.

V. CONCLUSION AND FUTURE WORK

This paper described how BPMN model examples are presented on the Web. Particularly, three categories of model elements (Activity, Event, and Gateway) and their labels were examined.

In the Web examples, for the labels of atomic Activities called Tasks, there is common consensus to follow the recommended Verb Object Style, since the majority of the label examples for model elements on the Web follow this style. For non-atomic Activities (i.e., normal Sub processes), two ways of labeling are preferred: Nominalization of a verb (Action Noun Style) and the recommended Verb Object style.

In the case of Event types, it turned out that the labeling styles vary depending on the Event type used. Additionally, within the same Event type, variations of labeling styles exist. For some of these labeling strategies, existing literature would even state that these labels have deficiencies. Therefore, it would be good that both providers of such examples and readers of these examples have a more critical look on them. However, it also must be said that some of the label examples for Events (e.g., Timer Event) make sense with respect to the certain type of Event.

If Gateways have a label, then it is quite well understood that a question mark (“?”) should close the label, as it is suggested in literature. But, this is the only accordance with literature. Since many of the Gateways do not even have any label, this could be also interpreted that labeling of Gateways for a better understanding of the process paths is not yet understood as an important feature by the community who posts process model examples on the Web. In future, the label quality in other BPMN diagrams, which are provided on the Web, will be examined too.

REFERENCES


