Evaluation of an LLM-based Process Modeling Tool

First of all, thank you very much for taking the time to contribute to this research on the conversational and LLM-based generation of Process Models. In the following, you will be asked to assess and interactively design models with an LLM-based tool. Please note that the generation process is restricted to using only a limited subset of BPMN constructs. Therefore, please do your assessments relative to that subset:

- Tasks
- XOR-Blocks, Parallel-Blocks, and Event-based Gateways
- Do-While Loop Blocks
- Data Objects, Data Stores, and Data-flows
- Pools and Lanes
- Message Throw- and Catching Events, and Message-Flows
- Timer Events
- Start- and End Events

⚠ *Please note, data flows are currently rendered in the form of textual annotations. Please do not take this non-standard behavior into consideration for your assessments.*

Structure of the Survey

Part 1: Demographics Questions

Part 2: Initial Model Evaluation

- You will be shown models that were generated based on two different process descriptions.
- You are then asked to assess the perceived quality of the models relative to the provided textual description.

Part 3: Tool Usage

- You will use the tool to complete already started modeling conversations.
- During the task, you can invite modeling experts in the form of model checkers to the chat.
- After you have successfully finished each modeling task, you will be asked for feedback on the experience for that model.

Progress and Breaks

Youcan take breaks between tasks. Your progress is saved after the submission of each task, so you can continue where youleft off.

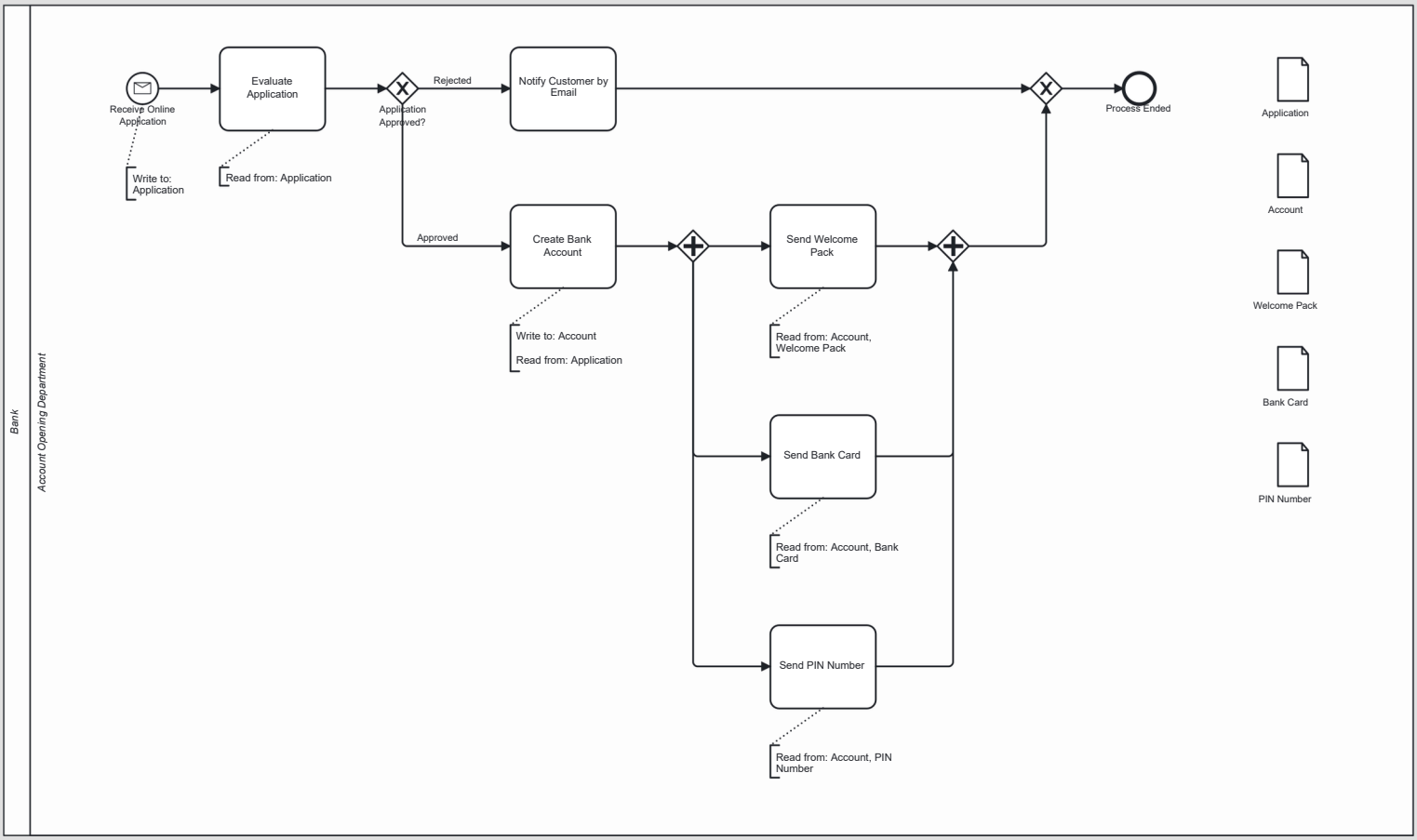
⚠ *Once You submit a task, you can not go back to it. If you still need to edit previous answers, please contact us via email.*

Start

M1.1 - Initial Model Evaluation

Textual Description

When the bank receives a new online application for opening a bank account, the application is evaluated. If the application is rejected, the customer is notified by email and the process ends. If the application is approved, a new bank account is created. Before the process ends, the bank sends a welcome pack, a bank card, and a PIN number in separate letters to the customer.



Tip: If the image is unclear, click here to open the image in a new tab to view it in full size.

Note: The data flows are represented in the form of comments. Please judge the modeled data-flow independent of its representation.

Model Evaluation Instructions

Pleaseevaluatethe modelsrelativetoonlythesupported BPMN constructs.

▶ Supported BPMN constructs

Please assess the quality of the generated process model based on the following criteria:

- Step 1: Provide your *overall* evaluation of the model.
- Step 2: Evaluate the model from *specific perspectives*: Pools & Lanes, Message-Flow, and Data-Flow.

1. Overall Correctness: The generated model represents the business process correctly.*

Fully Disagree (1) (2) (3) (4) (5) (6) (7) Fully Agree

2. Correctness: The generated model represents the business process correctly in terms of:

	1 (Fully Disagree)	2	3	4 (Neutral)	5	6	7 (Fully Agree)
Pools & Lanes							
Message-Flow							
Data & Data-Flow							

3. Overall Relevance: All the elements in the generated model are relevant for the representation of the business process.*

Fully Disagree 1 2 3 4 5 6 7 Fully Agree

4. Relevance: All the elements in the generated model are relevant for the representation of the business process in terms of:

	1 (Fully Disagree)	2	3	4 (Neutral)	5	6	7 (Fully Agree)
Pools & Lanes							
Message-Flow							
Data-Flow							

5. Overall Completeness: The generated model gives a complete representation of the business process.*

Fully Disagree 1 2 3 4 5 6 7 Fully Agree

6. Completeness: The generated model gives a complete representation of the business process in terms of:

	1 (Fully Disagree)	2	3	4 (Neutral)	5	6	7 (Fully Agree)
Pools & Lanes	○	○	○	○	○	○	○
Message-Flow	○	○	○	○	○	○	○
Data-Flow	○	○	○	○	○	○	○

7. Final Remarks - Comments on the Diagram:

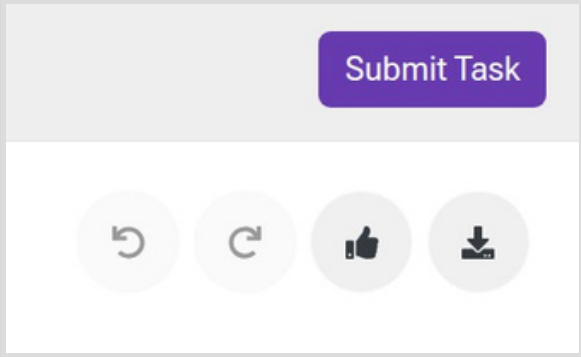
8. Final Remarks - Comments on the Questions:

Submit

Tool Usage Evaluation

In this part, you will use the tool to complete already started modeling conversations. You will be working on models that you have evaluated in the previous task.

Important Functionalities: You can navigate back to a previous model in the conversation and continue from it (especially useful if a comment has led to a worse model than before), run checkers, export the session, or submit the task when done.



Display of Models: You can zoom in and out in the usual way supported by your system: using Ctrl and +, Ctrl and –, Ctrl and your mouse wheel, or pinch to zoom.

In the conversation, can either:

- Directly provide feedback and ask the LLM for an improved version.
- Alternatively, invite modeling experts in the form of model checkers to the chat. Currently, there is one data-flow checker and one message-flow checker.

The model checkers can:

- Identify possible issues in your current model
- Provide suggestions for corrections
- Highlight elements relevant to the issue (Highlight)
- Allow you to automatically fix detected problems (AutoFix)

After you have successfully finished each modeling task, you will be asked for feedback on the experience for that model.

⚠ We encourage you to use the model checkers throughout the task to report in your feedback on their usefulness.
However, not every warning will identify a relevant issue for every scenario. It is up to you to decide if you want to resolve all the warnings and errors or accept the model regardless.

Start

M1.1 - Tool Usage Evaluation

In this follow-up, please evaluate the final model you generated using the tool. Your initial model evaluation answers are shown below.

1. Overall Correctness: The final model represents the business process correctly.*

Fully Disagree 1 2 3 4 5 **6** 7 Fully Agree

2. Overall Relevance: All the elements in the final model are relevant for the representation of the business process.*

Fully Disagree 1 2 3 4 5 **6** 7 Fully Agree

3. Overall Completeness: The final model gives a complete representation of the business process.*

Fully Disagree 1 2 3 4 **5** 6 7 Fully Agree

Tool Usage

4. The tool responded well to my feedback.*

Fully Disagree 1 2 3 4 5 Fully Agree

5. Did you use the Model Checkers?*

- ( ) Yes
- ( ) No

6. Did you use the Auto-Fix feature?*

- ( ) Yes
- ( ) No

7. It was frustrating to use the tool for the modeling task.*

Fully Disagree 1 2 3 4 5 Fully Agree

Submit

Final Tool Usage Evaluation

Perceived Usefulness and Intention to Use

Please rate your agreement with the following statements based on your experience with the anonymized tool name.

1. I think the tool makes modeling business processes easier.*

Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

2. I think the tool increases productivity in modeling business processes.*

Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

3. I think the tool is useful for modeling business processes.*

Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

4. If I could use the tool for my next modeling task, I would. *

Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

5. Do you have any general feedback or comments?

Submit