



universität
uulm

**Fakultät für
Ingenieurwissenschaften,
Informatik und
Psychologie**
Institut für Datenbanken
und Informationssysteme (DBIS)

Identifying GDPR-Critical Tasks in Business Processes using Large Language Models

Abschlussarbeit an der Universität Ulm

Vorgelegt von:

Merten Dieckmann
merten.dieckmann@uni-ulm.de
1058340

Gutachter:

Prof. Dr. Manfred Reichert
Prof. Dr. Rüdiger Pryss

Betreuer:

Magdalena von Schwerin

2025

Fassung 8. September 2025

© 2025 Merten Dieckmann

Satz: PDF- \LaTeX 2 _{ε}

Inhaltsverzeichnis

Abkürzungen	iv
1 Grundlagen von BPMN 2.0	1
1.1 Aktivitäten	2
1.2 Datenobjekte und Assoziationen	2
1.3 BPMN 2.0 XML	3
2 Einleitung	4
3 Theoretischer Hintergrund	5
4 Kriterien für die Modellauswahl	6
5 Konzept & Methodik	7
6 Datensatz & Testdesign	8
7 Evaluationsexperimente	9
8 Ergebnisse & Diskussion	10
9 Fazit & Ausblick	11
A Quelltexte	12
Literatur	13

Abkürzungen

LLM Large Language Model

BPMN Business Process Model and Notation

1 Grundlagen von BPMN 2.0

Dieses Kapitel thematisiert die für diese Arbeit notwendigen Grundlagen von BPMN 2.0.

Business Process Model and Notation (BPMN) ist eine Standard-Notation, um Geschäftsprozesse zu modellieren, welche von allen Teilnehmern verstanden werden können. Die aktuelle Version 2.0 des Standards wurde 2011 von der Object Management Group (OMG) entwickelt. Prozesse bestehen aus einer Abfolge von Flow-Elementen. Diese Flow-Elemente sind durch Sequenzflüsse miteinander verbunden, welche den Ablauf des Prozesses vorgeben. Der Umfang von Prozessen kann von Aufgaben für eine Person hin zu firmenweiten Tätigkeiten reichen [2].

Ein Ziel der OMG ist es, eine Brücke zwischen der Modellierung und der Implementierung zu schaffen. Der Standard kann im Dialog zwischen Geschäftspartnern jeder Ebene als schnelle Anforderungsübersicht dienen, den Entwicklern helfen, die Geschäftsprozessen zu implementieren und im Anschluss das Überwachen und Warten erleichtern. Ein anderes Ziel von BPMN 2.0 ist es, Prozesse visualisieren zu können, welche durch eine zur Ausführung entworfene XML Sprache modelliert sind [2].

In der BPMN 2.0 Spezifikation werden drei Unterprozesstypen beschrieben: Konversationsdiagramme, Kollaborationsdiagramme und Prozessdiagramme. Konversationsdiagramme modellieren die Logik vom Nachrichtenaustausch. Kollaborationsdiagramme beschreiben die Interaktion von verschiedenen Prozessteilnehmern. Prozessdiagramme sind für den Prozessablauf von Aktivitäten da und werden in der BPMN als Graph dargestellt [2]. In dieser Arbeit wird mit Prozessdiagrammen gearbeitet. Nachfolgend folgt eine Erklärung aller für diese Arbeit relevanten Flow-Elemente.

1.1 Aktivitäten

Eine Aktivität ist eine Tätigkeit, welche innerhalb von einem Geschäftsprozess ausgeführt wird. Es gibt drei Arten von Aktivitäten: Aufgabe, Subprozess und das Aufrufen von einer anderen Aktivität. Eine Aufgabe ist eine kleinste mögliche atomare Aktivität, welche meist von einem Benutzer ausgeführt wird. Sie wird durch ein abgerundetes Rechteck abgebildet (siehe Abbildung 1.1). Eine Aufgabe kann zudem einen Typen zugewiesen bekommen. Dieser wird oben links im Rechteck durch ein passendes Symbol dargestellt. Diese Eigenschaft wird im weiteren Verlauf der Arbeit für die Gamifizierung noch interessant werden [2].



Abbildung 1.1: Aufgabe ohne und Aufgabe mit Typ

1.2 Datenobjekte und Assoziationen

Mithilfe von Datenobjekten können Informationen im Laufe des Prozesses geladen und gespeichert werden. Assoziationen ermöglichen es, Informationen zwischen Datenobjekten und Aktivitäten zu transportieren. Sie haben keinen Einfluss auf den Prozessablauf. Datenobjekte werden durch eine Notiz und Assoziationen durch einen gestrichelten Pfeil mit offener Spitze dargestellt (siehe Abbildung 1.2). Die Informationen werden in Richtung des Pfeils transportiert [2].

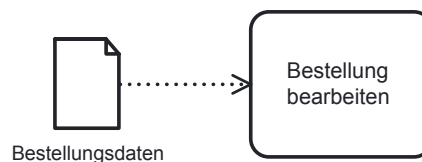


Abbildung 1.2: Datenobjekt mit Assoziation zu einer Aufgabe

1.3 BPMN 2.0 XML

BPMN 2.0 XML ist eine Auszeichnungssprache, welche auf einer von OMG erstellten XML Schema Definition (XSD) basiert [1]. Sie dient dazu, BPMN 2.0 Modelle abzuspeichern und somit einen Austausch zwischen unterschiedlichen Anwendungen zu ermöglichen. Eine BPMN 2.0 XML Datei hat die Endung *.bpmn* [2].

2 Einleitung

- Kontext & Motivation: BPMN-basierte Geschäftsprozesse, Bedarf an DSGVO-Konformität, Chancen/Risiken von Large Language Model (LLM).
- Problemstellung & Zielsetzung
- Forschungsfragen und Hypothesen.
- Aufbau der Arbeit: kurzer Überblick über die folgenden Kapitel.

3 Theoretischer Hintergrund

- Business Process Model and Notation (BPMN): Notation, modellgetriebene Analyse von Geschäftsprozessen.
- **Datenschutzrecht:** GDPR erklären.
- **Large Language Models:** Architektur, Training, Grenzen, Open-Source-Modelle & Datenschutz.
- Forschungslücken und aktueller Stand zur LLM-basierten Prozessanalyse.

4 Kriterien für die Modellauswahl

Siehe Excel

5 Konzept & Methodik

- Software-Architektur (Gesamtübersicht, Modul- und Datenflussdiagramme).
- Datenaufbereitung: BPMN-Parsing, Tokenisierung, Anonymisierung.
- Klassifikations-Pipeline: Prompt-Design bzw. Fine-Tuning, Label-Schema für GDPR-Kritikalität.
- Implementierungsdetails: genutzte Frameworks, Deployment (Docker u. a.), ...

6 Datensatz & Testdesign

- Generierung bzw. Sammlung repräsentativer BPMN-Modelle.
- Label-Prozess, Qualitätskontrolle
- Daten-Splits (Train/Validation/Test), falls Finetuning im Spiel ist

7 Evaluationsexperimente

- Versuchsaufbau
- Vergleichsbaselines und alternative Ansätze.
- Metriken (Precision, Recall, F1, AUC, Confidence-Scores) & Signifikanztests.
- Automatisiertes Testframework, Reproduzierbarkeit der Experimente.
- Fehlerquellen und Gegenmaßnahmen.

8 Ergebnisse & Diskussion

- Quantitative Resultate (Tabellen, Diagramme, Statistik).
- Qualitative Analyse
- Vergleich der Modelle & Ansätze; Interpretation der Unterschiede.
- Limitationen der Studie und Einflussfaktoren.

9 Fazit & Ausblick

- Zusammenfassung der Hauptekenntnisse bezogen auf die Forschungsfragen.
- Praktische Implikationen für Unternehmen & Regulatoren.
- Empfehlungen für zukünftige Arbeiten (Modell-Verbesserungen, zusätzliche Datenquellen, Pilotprojekte).

A Quelltexte

In diesem Anhang sind einige wichtige Quelltexte aufgeführt.

```
1 public class Hello {  
2     public static void main(String[] args) {  
3         System.out.println("Hello World");  
4     }  
5 }
```


Literatur

- [1] OMG. *About the Business Process Model And Notation Specification Version 2.0*. 2011. URL: <https://www.omg.org/spec/BPMN/2.0/About-BPMN> (besucht am 20.04.2023).
- [2] OMG. *Business Process Model and Notation (BPMN)*. 2011. URL: <https://www.omg.org/spec/BPMN/2.0/PDF> (besucht am 01.05.2023).

Name: Merten Dieckmann

Matrikelnummer: 1058340

Erklärung

Ich erkläre, dass ich die Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Ulm, den

Merten Dieckmann