



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 19. Juni 2025

© 2025 Merten Dieckmann

Satz: PDF- $\text{\LaTeX}$  2 <sub>$\epsilon$</sub>

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Theoretischer Hintergrund</b>	<b>2</b>
<b>3</b>	<b>Kriterien für die Modellauswahl</b>	<b>3</b>
<b>4</b>	<b>Konzept &amp; Methodik</b>	<b>4</b>
<b>5</b>	<b>Datensatz &amp; Testdesign</b>	<b>5</b>
<b>6</b>	<b>Evaluationsexperimente</b>	<b>6</b>
<b>7</b>	<b>Ergebnisse &amp; Diskussion</b>	<b>7</b>
<b>8</b>	<b>Fazit &amp; Ausblick</b>	<b>8</b>
<b>A</b>	<b>Quelltexte</b>	<b>9</b>
	<b>Literatur</b>	<b>10</b>

# 1 Einleitung

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

## 2 Theoretischer Hintergrund

- **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.

# **3 Kriterien für die Modellauswahl**

Siehe Excel

## 4 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.), ...

## 5 Datensatz & Testdesign

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



## 6 Evaluationsexperimente

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

## 7 Ergebnisse & Diskussion

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

## 8 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] Jörg Knappen. *Schnell ans Ziel mit LATEX 2e*. 3., überarb. Aufl. München: Oldenbourg, 2009.
- [2] Frank Mittelbach, Michel Goossens und Johannes Braams. *Der Latex-Begleiter*. 2., überarb. und erw. Aufl. ST - Scientific tools. München [u.a.]: Pearson Studium, 2005.
- [3] Joachim Schlosser. *Wissenschaftliche Arbeiten schreiben mit LATEX : Leitfaden für Einsteiger*. 5., überarb. Aufl. Frechen: mitp, 2014.
- [4] Thomas F. Sturm. *LATEX : Einführung in das Textsatzsystem*. 9., unveränd. Aufl. RRZN-Handbuch. Hannover [u.a.]: Regionales Rechenzentrum für Niedersachsen, RRZN, 2012.
- [5] Herbert Voß. *LaTeX Referenz*. 2., überarb. u. erw. Aufl. Berlin: Lehmanns Media, 2010.

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