

# Automate Text Processing for Schematically Analyzing Legal Texts

Magnus Bender<sup>1</sup>

<sup>1</sup> *Institute of Humanities-Centered Artificial Intelligence (CHAI), Universität Hamburg, Warburgstraße 28, 20354 Hamburg, Germany*

## Abstract

Legal texts are regulations specified in a general and abstract way: For specific situations, they determine the (legal) consequences of the situation on the basis of certain conditions that must or must not be met. Large Language Models (LLMs) are powerful tools for analyzing texts, but are prone to hallucinations and do not provide an explanation for their answers. This paper proposes to apply LLMs on legal texts in a way to mitigate the issues of LLMs while utilizing their powerful abilities to process text. The amount of legal texts is quite huge and constantly changing. Thus, it is not reasonable to manually process and extract facts required by intelligent agents working with legal text. This paper describes a concept for automating text processing for schematically analyzing legal texts. We provide a simple legal case example, describe how to model legal texts schematically, and to extract these schemes using LLMs. We conclude with a short evaluation of the abilities of ChatGPT and Gemini in the field of legal texts. Generally, our approach is not limited to legal texts and can be used to extract schemes from natural language texts helping intelligent agents to improve their decisions.

## Keywords

Legal Texts, Scheme Extraction, Intelligent Agent, Automatic Text Processing, Subjective Content Descriptions (SCDs)

## 1. Introduction

Large Language Models (LLMs) [1] are a powerful tool for analyzing texts and are used to provide chat bots, e.g., ChatGPT<sup>1</sup> or Gemini<sup>2</sup>. However, LLMs are prone to hallucinations, erroneous answers invented by the LLM, and do not provide an explanation for their answers. At least, the explanations do not allow full traceability, i.e., do not contain all the reasons taken into account while creating an answer. In some applications, such unprovable answers may be sufficient, but in others they are not. The AI-Act of the European Union (EU) [2] classifies systems into four risk levels. The higher the risk level of a system is, the more strict the obligations of the system is. Hence, explainability and traceability while creating answers in an intelligent system is an important field of research.

Let us consider the example of legal texts: Laws regulate certain situations and determine, on the basis of specified conditions that must or must not be fulfilled, which (legal) consequences result from the situation. Thereby, laws are specified in a general and abstract way, which needs to be applied on each concrete case. Additionally, laws—especially when it comes to the application of law—follow a hierarchically structure, e.g., to determine if an act is unlawfully, other laws, which might justify the act, need to be checked. Altogether, the laws provide a legislation which builds a scheme of conditions and entailments—at least in a simplified perspective without considering all the difficulties interpreting specific (individual) cases and human factors.

In this paper, we do not focus on automating court orders, but in automating text processing for schematically analyzing legal texts. Having extracted such schemes allows intelligent agents [3] to improve their decisions and provide better explanations to their principals, especially under human

---

*Humanities-Centred AI (CHAI), 4th Workshop at the 47th German Conference on Artificial Intelligence, September 23, 2024, Würzburg, Germany*

✉ [magnus.bender@uni-hamburg.de](mailto:magnus.bender@uni-hamburg.de) (M. Bender)

🌐 <https://www.philosophie.uni-hamburg.de/chai/personen/bender.html> (M. Bender)

🆔 0000-0002-1854-225X (M. Bender)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

<sup>1</sup><https://chatgpt.com/>

<sup>2</sup><https://gemini.google.com/>

supervision and, e.g., integrated in a system like AutoGPT [4]. The amount of legal texts is quite huge and constantly changing, e.g., on 1. January 2022 there have been 1 773 federal laws within total 50 738 law paragraphs in Germany<sup>3</sup> not including EU and other legal regulations. In view of these numbers, it is not reasonable to manually extract legal schemes which can be used by intelligent agents. We argue that current LLMs have the ability semantically analyze legal texts and automatically extract the required schemes. The extracted schemes then represent the legislation which can be used by an agent and also described to a human, to generate substantiate answers.

Generally, legal texts are only one kind of texts which describe conditions entailing consequences. For example, a manual of a device or a software follows a similar structure and may be processed in a similar way to extract schemes. However, for manuals it less important to get a traceable and verifiable chain of decisions made by an agent while generating the answer to a question.

The remainder of this concept paper is structured as follows: First, we give a simple legal case example from the field of German criminal law. We introduce basics concepts and three required laws needed to solve the example in the old-fashioned human way. Second, we describe Subjective Content Descriptions (SCDs) [5, 6] and how to use them to model legal texts schematically. Afterwards, we describe how to use LLMs for analyzing legal texts and creating SCDs forming a scheme. Next, we evaluate ChatGPT and Gemini on our legal case example. Finally, we conclude with a summary and outlook.

## 2. Legal Case Assessments

In this section, we present a simple exemplary case in the field of German criminal law. We describe how the case is solved by applying the German Criminal Code (Strafgesetzbuch, StGB)<sup>4</sup> and introduce three law paragraphs required for the solution. The example is based on German jurisprudence and the German language version of the StGB. We translate it into English for this paper. Note: While the laws are real and the solution of the legal case is correct, the legal explanation is heavily simplified.

### 2.1. Exemplary Legal Case

We start by giving an overview of the case and introducing the solution. Afterwards, we present a special structure used by legal professionals to report solution of legal cases.

#### **Case**

Walker *W* is out for a walk in the city park. Suddenly, *W* realizes that thief *T* is trying to steal *W*'s laptop from his backpack. *W* does not want to put up with this and grabs *T* firmly by the wrist, causing *T* to suffer a bruise and flee.

*Criminal liability of W applying the German Criminal Code (Strafgesetzbuch, StGB).*

We are asked to determine the criminal liability of *W*. *W* grabbed *T* by the wrist and caused a bruise on *T*'s arm. This action may be a bodily harm defined in the following paragraph.

#### **§ 223 I StGB – Bodily Harm**

Anyone who physically abuses another person or harms their health is liable to a custodial sentence not exceeding five years or a monetary penalty.

§ 223 I StGB requires either physical abuse or harmed health. Physical abuse is defined as any nasty, inappropriate treatment that has a more than insignificant impact on physical well-being or physical integrity. Harmed health is defined as causing or increasing a pathological condition, i.e., a condition that deviates from the normal functioning of the body. *W* caused a bruise on *T*'s arm, a bruise is an impact on the physical integrity of *T* and thus, a bodily harm.

<sup>3</sup><https://www.bundestag.de/presse/hib/kurzmeldungen-882012>, <https://dserver.bundestag.de/btd/20/007/2000721.pdf>, last accessed 28. June 2024

<sup>4</sup><https://www.gesetze-im-internet.de/stgb/>, English translation [https://www.gesetze-im-internet.de/englisch\\_stgb/](https://www.gesetze-im-internet.de/englisch_stgb/)

Objectively considered, W has committed bodily harm, but that does not directly mean that T is criminal liable. There might be a justification like the following:

### **§ 32 StGB – Self-Defense**

- I Anyone who commits an act that is required by self-defense is not acting unlawfully.
- II Self-defense is the defense required to avert a present unlawful attack from oneself or another.

§ 32 StGB defines a justification, the self-defense. The first sentence (I) defines that self-defense is a justification while the second sentence (II) describes the conditions of self-defense. Thus, if W caused the bruise on T's arm in self-defense, W is not criminal liable for the bodily harm of T's arm. For having a self-defense situation against T, W must be in an unlawful attack of T. This attack might be T's trial to steal W's laptop.

### **§ 242 I StGB – Theft**

Anyone who takes another person's movable property with the intention of unlawfully appropriating the property for themselves or a third party is liable to a custodial sentence not exceeding five years or a monetary penalty.

Theft, defined in § 242 I StGB, requires to take another person's movable property with the intention unlawfully appropriating the property. The laptop is a movable property owned by W. T tries to take the laptop from W's backpack without consent of W. Thus, T commits an unlawful attack to W.

Summarized, W is a self-defense situation regarding T trying to steal W's laptop. The bodily harm of T's arm caused by W is justified by self-defense. Hence, W is not criminal liable.

## **2.2. Reports of Legal Cases**

Law professionals writing a report about a law case follow a strict structure while solving the case. For German criminal law, there are three steps with multiple sub-steps. Our exemplary case needs to be structured as follows:

### **A. Elements of Offense**

*Assure that W's act fulfils the elements of the offense, i.e., a bodily harm (§ 223 I StGB) of T's arm.*

#### **I. Objective Elements**

*Assure that the conditions of bodily harm are objectively met. Physical abuse or harmed health of T, bruise is physical abuse. Furthermore, the bruise needs to be causally and objectively attributed to W's act.*

#### **II. Subjective Elements**

*Subjectively is required that W intended his act, i.e., acted with knowledge and will regarding the objective elements.*

### **B. Unlawfulness**

*Check if any conditions of justifications like self-defense (§ 32 StGB) are met. The fulfillment of a justification cancels the unlawfulness of an act.*

#### **I. Self-Defense Situation**

*Assure an ongoing current unlawful attack. This is the theft (§ 242 I StGB) of T trying to steal W's laptop. The theft itself again requires to be checked by using the three steps 1. Elements of Offense, 2. Unlawfulness, 3. Liability including potential sub-steps.*

#### **II. Self-Defense Action**

*Check, if W's action is necessary, suitable, and appropriate. Causing a bruise as a defense against the theft of a laptop is lawful.*

### C. Liability

*Finally, each offender needs to be liable for their act.*

After considering a legal case including relevant laws and the solution including how to structure an assessment with the solution, we move to SCDs for schematically representing laws.

## 3. Subjective Content Descriptions

Before we describe how to use LLMs to analyze legal text schematically, we need to consider representing the extracted schemes using SCDs. A typical use case of SCDs is an Information Retrieval (IR) agent. Such an agent works with a corpus consisting of text documents and offers an IR service: Given a query consisting of a sequence of words, relevant text documents in the corpus, including the exact positions such as sentences and words, are returned. Furthermore, an IR agent must maintain its corpus and thus identify different categories of text documents, e.g., distinguish between new and known text documents [7].

In the context of legal texts, an IR agent should maintain a corpus of laws and other legal texts, e.g., legal commentaries and court decision. All the information in the corpus may be relevant for a legal professional working on a certain problem. Besides retrieving relevant documents and opinions from the documents in the maintained corpus [6], the IR agent needs to *understand* law and be able to entail legal consequences based on conditions. Having schemes of the law documents support the IR agent with its tasks.

Furthermore, SCDs can be used to trace back from generated text to locations in the source documents. For example, the Humanities Aligned Chatbot (ChatHA) [8] uses SCDs to add citations to LLM generated texts. Tracing back can be used to explain generated results.

SCDs are a good choice for legal texts, because they represent *subjective* descriptions and in the field of law, things are often controversial, i.e., different opinions prevail. SCDs assume persons having different perspectives on the text documents in a corpus, depending, among other things, which task a person wants to process with the help of the corpus and an IR agent. This results in different requirements for the IR agent and, in particular, different subjective perspectives on the same corpus. These different subjective perspectives must be taken into account and constitute the *subjectivity* of SCDs.

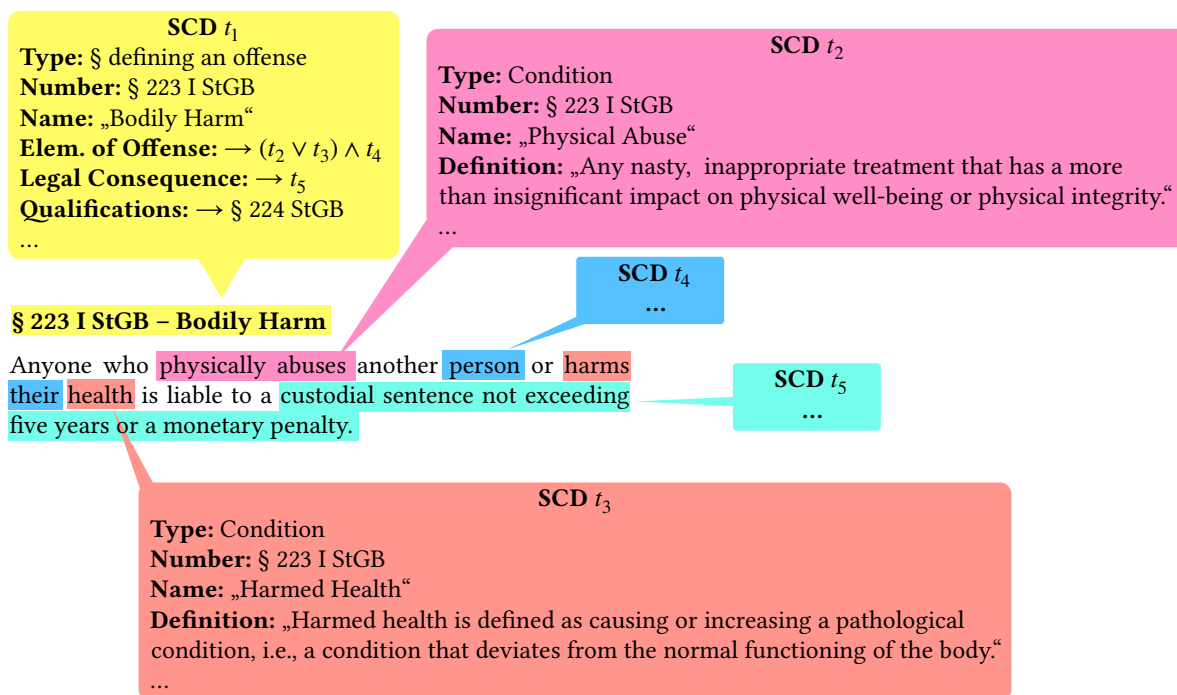
### 3.1. Formalizing SCDs

SCDs are content descriptions that are linked to locations, e.g., sentences or words, in a corpus of text documents. SCDs contain additional data for the understanding and automatic processing of natural language text. For example, an SCD (i) references several similar sentences in the agent’s corpus, (ii) contains additional hints for interpretation, or (iii) relations to, e.g., complementary SCDs [7]. Similar to sticky notes on a piece of paper, SCDs can be understood as hints that help an agent or human to understand a text document by providing descriptions, references, and explanations.

For a corpus  $\mathcal{D}$  of text documents  $d_1, \dots, d_M$  the SCD set  $g(\mathcal{D})$  contains all SCDs. Each SCD  $t_1, \dots, t_K \in g(\mathcal{D})$  includes referenced sentences and additional data. The referenced sentences are the locations in the corpus to which an SCD provides a description. The description themselves are stored in the additional data and are not limited to textual hints. Especially different kinds of relations to other SCDs are part of the additional data, too.

### 3.2. Modelling Law

For each law paragraph, multiple SCDs are used and each SCD contains different data. One SCD is associated with the entire paragraph and links with relations to the different elements this law paragraph requires for its offense. The different elements are then represented by an SCD which is located directly with the words in the law paragraph inducing this element. For each element the conditions to be met



**Figure 1:** Sketch of using SCDs to represent the scheme of § 223 I StGB using SCDs.

are listed including potential sub-elements, again as links to SCDs. Whereas the latter SCDs may be referencing a legal commentary.

Next, we give an explicit example of representing law with SCDs and how LLMs can be used to extract these representations.

## 4. Schematically Analyzing Legal Texts

This paper proposes an approach for extracting SCDs representing schemes from legal texts. First, we give an example of the representation and then show that the content of the representation can be extracted from an LLM, e.g., ChatGPT.

### 4.1. Representation

An example for representing § 223 I StGB using SCDs is sketched in Figure 1. The law paragraph is depicted in the middle of the figure. The locations of the SCDs, i.e., sequences of words, are highlighted in different colors, each color represents one SCDs. SCD  $t_1$  (yellow) is a general description of § 223 I StGB. It contains among others the elements of the offense as expression of SCDs to check and the legal consequence. Additionally, there is a link to another law paragraph or rather its SCDs which is a qualification, i.e., more strict penalty in certain cases, of § 223 I StGB. SCDs  $t_2$  (purple) and  $t_3$  (red) represent each one element of § 223 I StGB, including a definition of the element. This definition could imply further elements, which would be modelled by further SCDs.

With this representation, an IR agent checking § 223 I StGB gets the elements of the offense and the legal consequences. The agent furthermore gains the information about each element and the required sub-steps to check the different elements. Overall, this representation is a proposal which needs to be firmed up and formalized in detail. The three steps of 1. *Elements of Offense*, 2. *Unlawfulness*, 3. *Liability* would also be formalized by SCDs, as they are defined in text of the StGB, too.

## 4.2. Extraction by LLM

The different data to be stored in the SCDs associated with laws are extractable using an LLM, e.g., ChatGPT. We prepare the following question for ChatGPT and ask for a detailed list of conditions. Note, that we added the law itself and definitions how they are found in legal commentaries. This includes to add legal commentaries to the IR agent's corpus.

### § 223 I StGB – Bodily Harm

Anyone who physically abuses another person or harms their health is liable to a custodial sentence not exceeding five years or a monetary penalty.

#### Definitions

- Physical abuse is defined as any nasty, inappropriate treatment that has a more than insignificant impact on physical well-being or physical integrity.
- Harmed health is defined as causing or increasing a pathological condition, i.e., a condition that deviates from the normal functioning of the body.

*Schematically analyze the law and definition above and list explicitly the conditions required for acting a bodily harm.*

For the above question, we get the following answer from ChatGPT (version 4o). The displayed answer is an extract of ChatGPT's full answer.

1. Physical Act
  - Physical Abuse
    1. **Conduct** The perpetrator engages in conduct that involves the treatment of another person.
    2. **Nature of Conduct** The conduct is nasty or inappropriate.
    3. **Impact on Victim** The conduct has a more than insignificant impact on the victim's physical well-being or physical integrity.
  - Harmed Health
    1. **Pathological Condition** The conduct causes or increases a pathological condition.
    2. **Deviation from Normal Functioning** The pathological condition must deviate from the normal functioning of the body.
2. Mental State
  - **Intentional or Negligent** The perpetrator must have acted with either intent or negligence in causing the physical abuse or harm to health.

The answer demonstrates that ChatGPT is able to identify the two main elements "Physical Abuse" and "Harmed Health". In both cases it also correctly selects the definition and identifies the relevant parts, i.e., for "Physical Abuse": The 1. treatment of another person, which is 2. nasty or inappropriate, and 3. with more than insignificant impact to the well-being or physical integrity. With post-processing and more detailed questions, such a list can be used to create the required SCDs.

ChatGPT also demonstrates, that it already knows the StGB because it adds a "mental state" to its list of conditions. This "Mental State" represents the *subjective elements* required in German criminal law while "Physical Act" represents the *objective elements* (see Subsection 2.2). The demand of *subjective elements* in German criminal law also comes directly from the StGB.

## 5. Evaluation

We evaluate our exemplary case with the chat bots ChatGPT and Gemini and demonstrate that both are not able to create a full legal case assessment. We test three different models, i.e., Gemini version 2024.06.18, ChatGPT 3.5, and ChatGPT 4o.



For the evaluation, we use the case introduced in Subsection 2.1 in English language and use it verbatim as prompt. All three models are able to cite numbered paragraphs from the StGB, hence we do not need to add the laws of the StGB to the prompt.

## 5.1. Results

The answers of all three models do not follow the structure for legal case assessments. Between ChatGPT 3.5 and 4o is no big difference. Both ChatGPT models start with a description of self-defense (§ 32 StGB) and then correctly state that T's attempted theft is an unlawful attack, to which self-defense is considered lawful. ChatGPT 4o discusses the bodily harm in detail, while ChatGPT 3.5 just implicitly assumes it.

Gemini follows the structure for legal case assessments slightly better, it starts with discussing coercion. Coercion is a different offense in the StGB, however for our exemplary case, it's a very far-fetched idea. Next, Gemini correctly describes and assumes self-defense (§ 32 StGB).

Interestingly, no model mentions theft with § 242 I StGB. The word *theft* is used, but it seems that all models directly anticipate a theft without checking its conditions using § 242 I StGB.

Summarized, ChatGPT 4o does the best job, while all three models correctly state that W is not liable because of self-defense. However, no model follows the structure for legal case assessments and mentions all relevant laws of the StGB needed to properly solve the case.

## 5.2. Discussion

Current LLMs and chat bots have been trained on the StGB and are able to answer questions about it. However, they do not follow the structure for solving legal cases required by German law professionals. Additionally, for our simple exemplary case, they do not entirely check the case and omit checking for theft.

The results support our argumentation that current LLMs are able to semantically analyze legal texts, but need more enforcement to follow required schemes when working with law.

Next, we summarize our proposed ideas and give an outlook.

## 6. Conclusion

This paper proposes an approach for schematically analyzing legal texts. We start with introducing an exemplary case from the field of German criminal law and describe its solution. Next, we introduce SCDs and why SCDs are a perfect choice to model (legal) conditions and consequences in corpora of text documents for an IR agent. Afterwards, we combine law and SCDs: We describe the representation using SCDs with the exemplary case and how this representation can be generated with LLMs like ChatGPT. In the end, we do a short evaluation to demonstrate that currently ChatGPT and Gemini are not able to satisfyingly solve the exemplary case.

The proposed approach is not limited to legal texts and solving cases, but much more suitable for a general problem: It allows IR agents to extract and use text induced logical interdependencies. Especially, our approach is not intended to automate court orders, but supporting legal professionals during information gathering and research in legal fields.

In general, the approach has similarities to Retrieval Augmented Generation (RAG) [9], but the focus is not on providing more texts to a model, but on analyzing the available texts in more detail. Future work will focus on building the proposed system and step by step realizing the sketched ideas.

## References

- [1] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin, Attention is all you need (2017).

- [2] AI-Act, Proposal for a regulation of the european parliament and of the council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts, 2021. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52021PC0206>.
- [3] S. Russell, P. Norvig, *Artificial Intelligence, Global Edition A Modern Approach*, Pearson Deutschland, 2021.
- [4] H. Yang, S. Yue, Y. He, *Auto-gpt for online decision making: Benchmarks and additional opinions*, 2023. URL: <https://arxiv.org/pdf/2306.02224.pdf>.
- [5] F. Kuhr, M. Bender, T. Braun, R. Möller, *Augmenting and automating corpus enrichment*, *Int. J. Semantic Computing* 14 (2020) 173–197.
- [6] M. Bender, T. Braun, R. Möller, M. Gehrke, *Enhancement of subjective content descriptions by using human feedback*, 2024. URL: <https://arxiv.org/abs/2405.15786>.
- [7] M. Bender, F. Kuhr, T. Braun, *To extend or not to extend? Enriching a corpus with complementary and related documents*, *Int. Journal of Semantic Computing* 16 (2022) 521–545.
- [8] T. Asselborn, S. Melzer, S. Aljoumani, M. Bender, F. A. Marwitz, K. Hirschler, R. Möller, *Fine-tuning bert models on demand for information systems explained using training data from pre-modern arabic*, in: *Proceedings of the Workshop on Humanities-Centred Artificial Intelligence (CHAI 2023)*, *CEUR Workshop Proceedings*, 2023, pp. 38–51. URL: <https://ceur-ws.org/Vol-3580/paper5.pdf>.
- [9] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel, et al., *Retrieval-augmented generation for knowledge-intensive NLP tasks*, *Advances in Neural Information Processing Systems* 33 (2020) 9459–9474.