

Appendix: Prompting Strategies and Evaluation

This appendix describes all prompting strategies used in our JURIX 2025 paper *"From Court Decisions to Guiding Principles: Advancing Complex Legal Summarization with LLMs"*. It includes full prompt examples as well as the evaluation prompt used for assessing guiding principle (*Leitsatz*) generation.

A. Prompting Strategies

Prompt design plays a pivotal role in shaping large language model (LLM) outputs. To assess the spectrum of prompting effectiveness, we implemented and compared four categories of prompting strategies.

A.1. Simple Prompts (Baselines)

As baselines, we designed three simple prompts reflecting how non-expert users might typically instruct an LLM:

SimplePrompt-v1:

```
You are a legal assistant.
Write a guiding principle (headnote) for the following court decision.

Court Decision:

    <<court decision>>

Output Format should be the following (only generated guiding principle
in German):
Guiding principle:
```

SimplePrompt-v2:

```
Read the following German court decision and formulate a Leitsatz that
reflects the central legal principle.

Court Decision:

    <<court decision>>

Output Format should be the following (only generated Leitsatz):
Leitsatz:
```

SimplePrompt-v3:

```
You are a legal expert preparing a Leitsatz for a case law database.
From the court decision below, draft a Leitsatz (headnote) summarizing
the key legal principle derived from the ruling.

    <<court decision>>

Output Format should be the following (only generated Leitsatz):
Leitsatz:
```

A.2. Expert-Structured Prompts

We adapted the structured design proposed by Steffes et al. [1], referred to as *ExpertStructuredPrompt*, which instructs the LLM to identify unresolved legal questions and formulate guiding principles in a structured manner. While it represents the most advanced legal expert-crafted prompt available, it has two limitations: (1) its question–answer format diverges from the concise, declarative style of guiding principles, and (2) it often generates additional or overly lengthy outputs. To address these limitations, we developed two refinements:

- *UnlabeledStructuredPrompt*, which removes the Q&A structure to align more closely with expert drafting style, and
- *UnlabeledStructuredPromptCoI*, which reformulates the prompt into a chain-of-instructions (CoI) [2] format to encourage conciseness and legal fidelity.

ExpertStructuredPrompt

Write the guiding principles for the following judgment. The guiding principles should reflect the key legal issues decided by the court, including the specific decision made by the court. Each legal question should be presented in a separate guiding principle. First state the legal question and then the corresponding guiding principle. Only state legal questions that were previously unanswered in the literature and were newly decided in the judgment.

Use the following template:

****RF****: Legal question

****LS****: Guiding principle

Your output should be in the original language of the judgment (i.e., German).

#####

Judgment:

<<court decision>>

UnlabeledStructuredPrompt

Write the guiding principles for the following judgment.

The guiding principles should reflect the key legal issues decided by the court, including the specific decision made by the court.

If the judgment addresses multiple legal questions, you should write one guiding principle per question.

However, please only consider legal questions that were previously unanswered in the literature and were newly decided in the judgment.

Your output should be in the original language of the judgment (i.e., German).

Only generate the guiding principles. Do not include any explanations. Present them without labels, subheadings, or captions. Use the format employed in German court decisions.

Judgment is here:

<<court decision>>

UnlabeledStructuredPromptCoI

Task: Write the guiding principles for the following judgment.

Instructions:

1. Read the judgement carefully.
1. Extract key legal issue(s) decided by the court, including the specific decision made by the court.
2. If the judgment addresses multiple legal issues, you should write one guiding principle per issue (only for the legal issue that was previously unanswered in the literature and were newly decided in the judgment).

Your output should be in the original language of the judgment (i.e., German).

Only generate the guiding principles. Do not include any explanations. Present them without labels, subheadings, or captions. Use the format employed in German court decisions.

Judgment is here:

<<court decision>>

A.3. *ChatCoI* (CoI Prompt Collaboratively Created by Human and LLM)

ChatCoI is a structured prompt composed of a chain of instructions (CoI) [2], collaboratively created by a human and a large language model (LLM) through interactive dialogue. Its purpose is to guide the LLM in constructing a prompt that can later be used to perform a specific task. The process began with a confirmation of the model's understanding of the concept of a *Leitsatz*, followed by a request for the model to propose step-by-step instructions that another LLM could follow to generate a *Leitsatz* from a court decision. An example of this interaction is shown below:

Interactive Dialogue for *ChatCoI* Prompt Creation

Human: What does “*Leitsatz*” mean in German court decisions?

LLM (ChatGPT-4o): [Explanation of the concept of a *Leitsatz*.]

Human: Yes, your understanding is correct. Now, could you suggest a chain-of-instructions (CoI) prompt for an LLM to automatically generate a *Leitsatz* based on a given judgment?

LLM (ChatGPT-4o): [Proposed CoI instructions, e.g.:

To generate a *Leitsatz*, you need to:

1. Understand the legal context – you must know what legal issue is being decided (e.g., remuneration, liability, official duties).
2. Identify the court's conclusion – what did the court ultimately decide about that issue?
3. Extract the core legal principle – summarize that decision as a rule or principle in neutral legal language.]

Due to space limitations, we only reproduced the shortest version of *ChatCoI* (*ChatCoI-v1*) in our JURIX paper, which is presented in the case box above. Using the same interactive dialogue in separate new chats, we also generated two additional variants: *ChatCoI-v2* and *ChatCoI-v3*. These longer versions contained more detailed and elaborate step-by-step instructions.

ChatCoI-v1

To generate a Leitsatz, you need to:

1. Understand the legal context - you must know what legal issue is being decided (e.g., remonstration, liability, official duties).
2. Identify the court's conclusion - what did the court ultimately decide about that issue?
3. Extract the core legal principle - summarize that decision as a rule or principle in neutral legal language.

Please generate a Leitsatz from the following court decision:

<<court decision>>

Output Format should be the following (only generated Leitsatz):

Leitsatz:

ChatCoI-v2

You are a legal assistant trained in German jurisprudence. Your task is to extract and formulate a Leitsatz (guiding principle) from a full court judgment. Follow the steps carefully.

◆ Step 1: Understand the Case

- * Read the full text of the judgment carefully.
- * Identify the main legal issue(s) the court resolved.
- * Note any specific statutes, legal doctrines, or precedent that play a central role.

◆ Step 2: Identify the Holding (Kernaussage)

- * Focus on the court's core legal finding.
- * Ask: What legal principle does the court apply or establish to resolve the main issue?
- * Make sure the holding is generalizable, i.e., it could apply in similar future cases.

◆ Step 3: Abstract the Rule

- * Reformulate the holding into a general legal proposition, avoiding case-specific facts.
- * If possible, include legal terminology or references to statutes or doctrines (e.g., § 280 BGB), but avoid unnecessary jargon.

◆ Step 4: Compose the Leitsatz

- * Write the Leitsatz in clear, formal, and concise German (if the output is to be in German).
- * Use passive voice or neutral phrasing commonly used in Leitsätze.
- * Avoid subjective language; be objective and authoritative.

◆ Step 5: Review and Refine

- * Ensure that the Leitsatz accurately reflects the court's legal reasoning without paraphrasing the entire judgment.
- * Confirm that it is legally precise, succinct (1-3 sentences), and not overly fact-dependent.

Please generate a Leitsatz from the following court decision:

<<court decision>>

✓ Output Format:

Leitsatz:

[Insert the generated Leitsatz here.]

ChatCoI-v3

Goal: Generate a legally accurate and concise Leitsatz that captures the core legal principle(s) established in the judgment.

Instructions:

1. Understand the Context
 - * Read the entire judgment carefully.
 - * Identify the legal area (e.g., civil law, criminal law, constitutional law).
 - * Note the key facts, legal issues, and applicable statutory provisions.
2. Extract the Legal Question(s)
 - * Identify the central legal question(s) the court had to answer.
 - * Focus on the most important legal issue that affected the outcome.
3. Analyze the Reasoning
 - * Summarize how the court interpreted the relevant laws or precedents.
 - * Note any novel interpretations, clarifications, or exceptions made by the court.
4. Formulate the Holding
 - * Convert the court's decision on the main legal issue into a general legal principle.
 - * Avoid case-specific details unless they define the rule's scope.
5. Write the Leitsatz
 - * Phrase the Leitsatz in formal, impersonal German legal language (or translate into English if required).
 - * Use the present tense and normative style.
 - * The Leitsatz should:
 - * Stand on its own, without needing the full judgment for context.
 - * Be concise, typically one or two sentences.
 - * Reflect the normative outcome of the case.
6. (Optional) Quality Check
 - * Ensure legal clarity, accuracy, and general applicability.
 - * Confirm that it aligns with the reasoning and outcome of the case.

Please generate a Leitsatz from the following court decision:

<<court decision>>

Output Format should be the following (only generated Leitsatz):
Leitsatz:

In addition, we created a few-shot variant, *ChatCoI-3Examples*. After confirming the model's understanding of the concept (by asking, "*What does Leitsatz mean in German court decisions?*"), we inserted an additional instruction:

"Yes, your understanding is correct. I will now provide you with three examples of judgments, each accompanied by a corresponding Leitsatz. Please learn from these examples how to construct a Leitsatz."

The LLM (ChatGPT-4o) then requested examples, to which we supplied three randomly chosen judgment–Leitsatz pairs. After providing these examples, we added another instruction:

"Now, could you suggest a chain-of-instructions (CoI) prompt for an LLM to automatically generate a Leitsatz based on a given judgment?"

The LLM then produced *ChatCoI-3Examples* as follows:

ChatCoI-3Examples

You are a legal expert specialized in German jurisprudence and legal drafting.

You will be given the full text of a German court judgment. Your task is to extract and formulate a precise **Leitsatz** (headnote) that summarizes the core legal principle established by the court in this case.

Follow these steps carefully:

1. **Read the Judgment Thoroughly**: Review the full reasoning of the decision. Focus on the key legal rule(s) that were decisive for the outcome.
2. **Identify the Normative Legal Principle**:
 - What legal issue did the court resolve?
 - What abstract rule does the decision establish for future, similar cases?
 - If multiple legal rules are central, determine if they can be expressed clearly in one or two sentences.
3. **Generalize the Facts**:
 - Remove all references to party names, dates, and procedural history.
 - Replace them with abstract categories (e.g., "a credit institution", "a shareholder", "an advertisement").
4. **Formulate the Leitsatz**:
 - Write in **formal, abstract, and normative language**.
 - Use declarative sentences (e.g., "Eine Klausel ist unwirksam, wenn..." or "Wird eine Forderung über... gestundet, handelt es sich um...").
 - Avoid case-specific details.
 - If needed, cite the **relevant legal norms** (e.g., § 307 Abs. 1 BGB), but do not quote full statutory text.
 - Use **no more than 2-3 sentences**.
 - If appropriate, include an introductory reference to precedent (e.g., "Anschluss an BGH...").
5. **Double-check for Precision**:
 - Ensure the Leitsatz captures the **holding**, not mere observations or dicta.
 - Make sure it can stand alone and make sense outside the judgment.

Output Format:

Leitsatz:

<Insert the resulting Leitsatz here>

A.4. *SelfCoI (Self-Generated Chain-of-Instructions)*

We additionally examined *SelfCoI*, a setting in which the LLM autonomously generated its own task-specific instructions during inference. In this setup, the LLM was first prompted to design an internal “chain-of-instructions” for itself before generating the final *Leitsatz*.

Prompt Given to the LLM:

```
What does "Leitsatz" mean in a German court decision?  
Could you create a Chain-of-Instructions (CoI) prompt to instruct GPT-4o  
to generate a Leitsatz for a given court decision?  
Then, please use the prompt you created to generate a Leitsatz.
```

```
Here is Court Decision:
```

```
<<court decision>>
```

```
Please do not show your created CoI prompt.  
Your final response must be formatted exactly like this:  
Leitsatz: <only your Leitsatz in German>
```

Because *SelfCoI* relies on internal prompt generation, its outputs may vary across runs. To account for this variability, we executed the *SelfCoI* procedure three times under identical conditions. The resulting outputs were denoted *SelfCoI-v1*, *SelfCoI-v2*, and *SelfCoI-v3*. This setup allowed us to evaluate the consistency of the approach and to determine whether autonomous prompt creation provides stable advantages over collaborative (*ChatCoI*) or expert-created domain-specific prompting strategies.

A.5. *LLM-Based Evaluation*

For evaluation, we used widely adopted metrics: BLEU, ROUGE, and METEOR for lexical overlap; and BERTScore, SBERT, and GPT embeddings (*text-embedding-ada-002*) for semantic similarity.

To better approximate expert judgment, we employed GPT-4o as a meta-evaluator. The model was instructed to act as a German legal writing expert and to evaluate candidate guiding principles against their reference according to three criteria: semantic fidelity, completeness, and conciseness — each rated on a 0–5 scale. It also assigned an overall score (0–100), reflecting a holistic quality judgment rather than a sum of the sub-scores. Following our design, an overall score of 50 was set as the acceptability threshold: candidates scoring below 50 were treated as inadequate guiding principles, whereas those scoring 50 or above were considered potentially usable in practice. This threshold was chosen to balance leniency in stylistic variation with strictness in semantic correctness.

We used GPT-4o to evaluate the generated *Leitsätze* based on the following prompt:

You are an expert in German legal writing.
Please evaluate how well the candidate Leitsatz matches the reference Leitsatz according to three criteria:

1. Semantic Fidelity (0-5): Does the candidate accurately capture the meaning and core legal reasoning of the reference?
2. Completeness (0-5): Does the candidate include all legally relevant aspects, without omitting essential elements?
3. Conciseness (0-5): Is the candidate expressed in a compact, stylistically appropriate manner, avoiding redundancy or unnecessary detail?

Output Format:

1. Semantic Fidelity (0-5):
2. Completeness (0-5):
3. Conciseness (0-5):

Overall Score (0-100):

Only output the scores.

Reference Leitsatz:

<<reference Leitsatz>>

Candidate Leitsatz:

<<candidate leitsatz>>

References

- [1] Steffes B, Wiedemann NT. Generating Guiding Principles: Evaluating Large Language Models for Complex German Legal Summaries. In JSAI International Symposium on Artificial Intelligence 2025 May 23 (pp. 49-65). Singapore: Springer Nature Singapore.
- [2] Zin MM, Satoh K, Borges G. Leveraging LLM for Identification and Extraction of Normative Statements. In Legal Knowledge and Information Systems 2024 (pp. 215-225). IOS Press.