

Controlling the Evolution of a Legal Knowledge Base for an In-Context Learning AI System - Design and Implementation

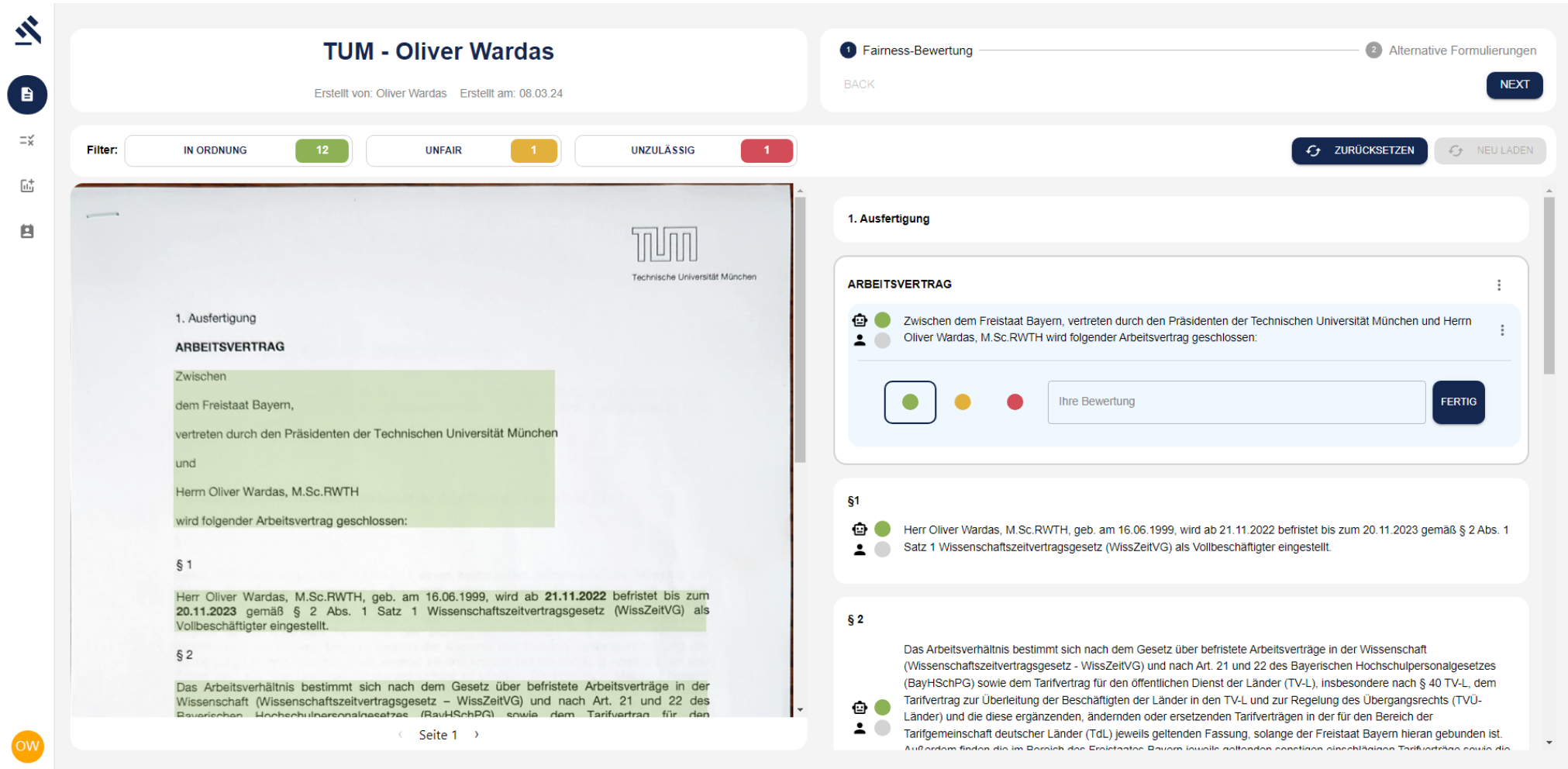
John Graf von Hardenberg

Sep 9, 2024, Bachelor's Thesis Kick-off Presentation

Chair of Software Engineering for Business Information Systems (sebis)
Department of Computer Science
School of Computation, Information and Technology (CIT)
Technical University of Munich (TUM)
www.matthes.in.tum.de

Background & Motivation

AI-Assisted Legal Analysis and Correction of German Employment Contracts



TUM - Oliver Wardas
Erstellt von: Oliver Wardas | Erstellt am: 08.03.24

Filter: **IN ORDUNG** 12 **UNFAIR** 1 **UNZULÄSSIG** 1 ZURÜCKSETZEN NEU LADEN

1. Ausfertigung

ARBEITSVERTRAG

Zwischen dem Freistaat Bayern, vertreten durch den Präsidenten der Technischen Universität München und Herrn Oliver Wardas, M.Sc.RWTH wird folgender Arbeitsvertrag geschlossen:

§ 1
Herr Oliver Wardas, M.Sc.RWTH, geb. am 16.06.1999, wird ab 21.11.2022 befristet bis zum 20.11.2023 gemäß § 2 Abs. 1 Satz 1 Wissenschaftszeitvertragsgesetz (WissZeitVG) als Vollbeschäftigter eingestellt.

§ 2
Das Arbeitsverhältnis bestimmt sich nach dem Gesetz über befristete Arbeitsverträge in der Wissenschaft (Wissenschaftszeitvertragsgesetz - WissZeitVG) und nach Art. 21 und 22 des Bayerischen Hochschulpersonalgesetzes (BayHSPG) sowie dem Tarifvertrag für den öffentlichen Dienst der Länder (TV-L), insbesondere nach § 40 TV-L, dem Tarifvertrag zur Überleitung der Beschäftigten der Länder in den TV-L und zur Regelung des Übergangsrechts (TVÜ-Länder) und die diese ergänzenden, ändernden oder ersetzenden Tarifverträgen in der für den Bereich der Tarifgemeinschaft deutscher Länder (TdL) jeweils geltenden Fassung, solange der Freistaat Bayern hieran gebunden ist. Außerdem finden die im Bereich des Freistaates Bayern jeweils geltenden sonstigen einschlägigen Tarifverträge sowie die

1. Fairness-Bewertung **2. Alternative Formulierungen**
BACK NEXT

Ihre Bewertung FERTIG

Seite 1

Background & Motivation

AI-Assisted Legal Analysis and Correction of German Employment Contracts

What is a Prüfanleitung?

Umfang von über **8** Stunden pro Woche bzw. als abgegoltene Überstunden ist bei einer Tätigkeit von 40 Stunden in der Woche unzulässig

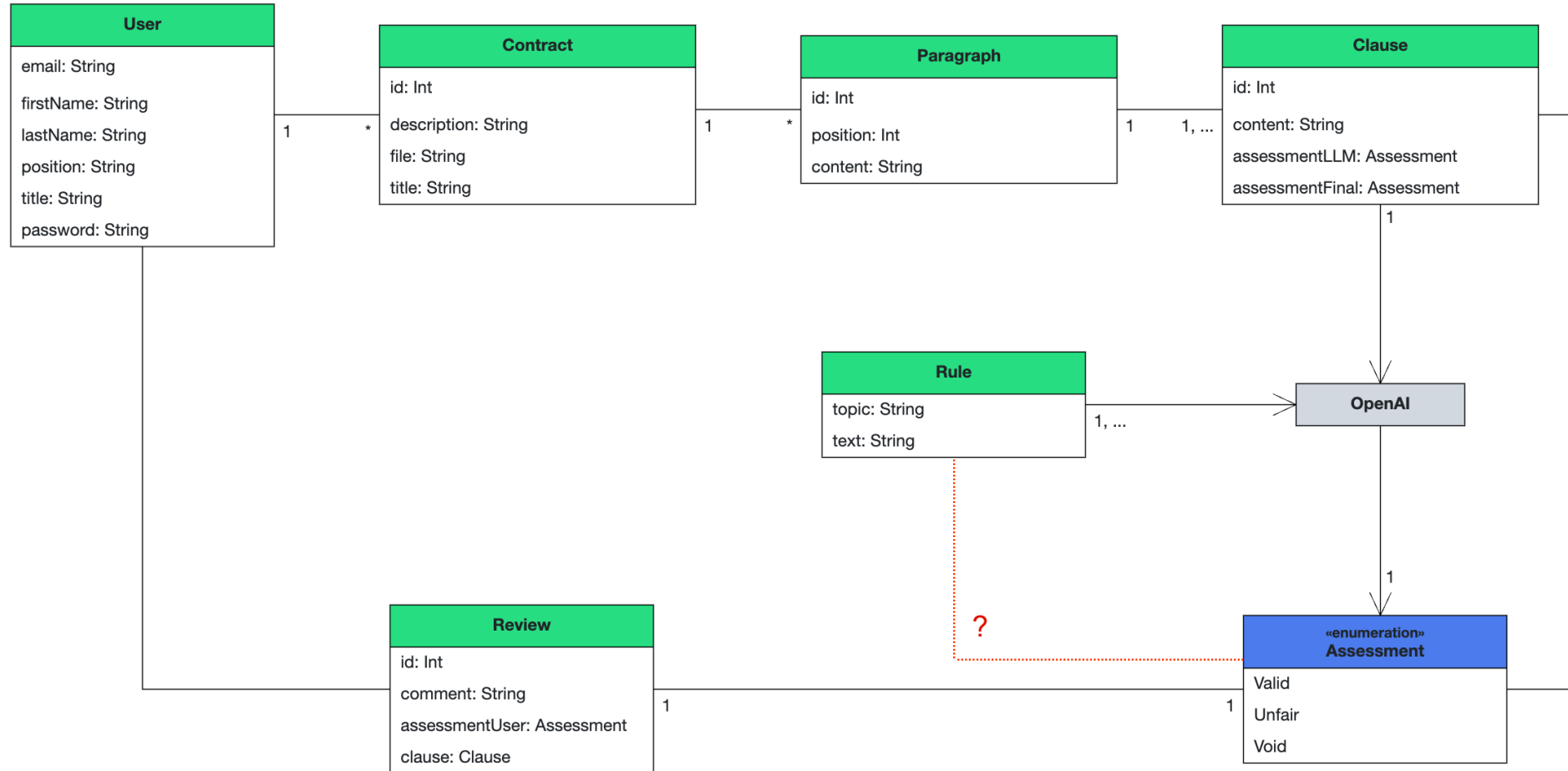
Change in law

Umfang von über **10** Stunden pro Woche bzw. als abgegoltene Überstunden ist bei einer Tätigkeit von 40 Stunden in der Woche unzulässig

- Law can change and therefore Prüfanleitungen can change
- Currently not possible to track the influence of these changes in the system
- What clauses have to be revalidated?

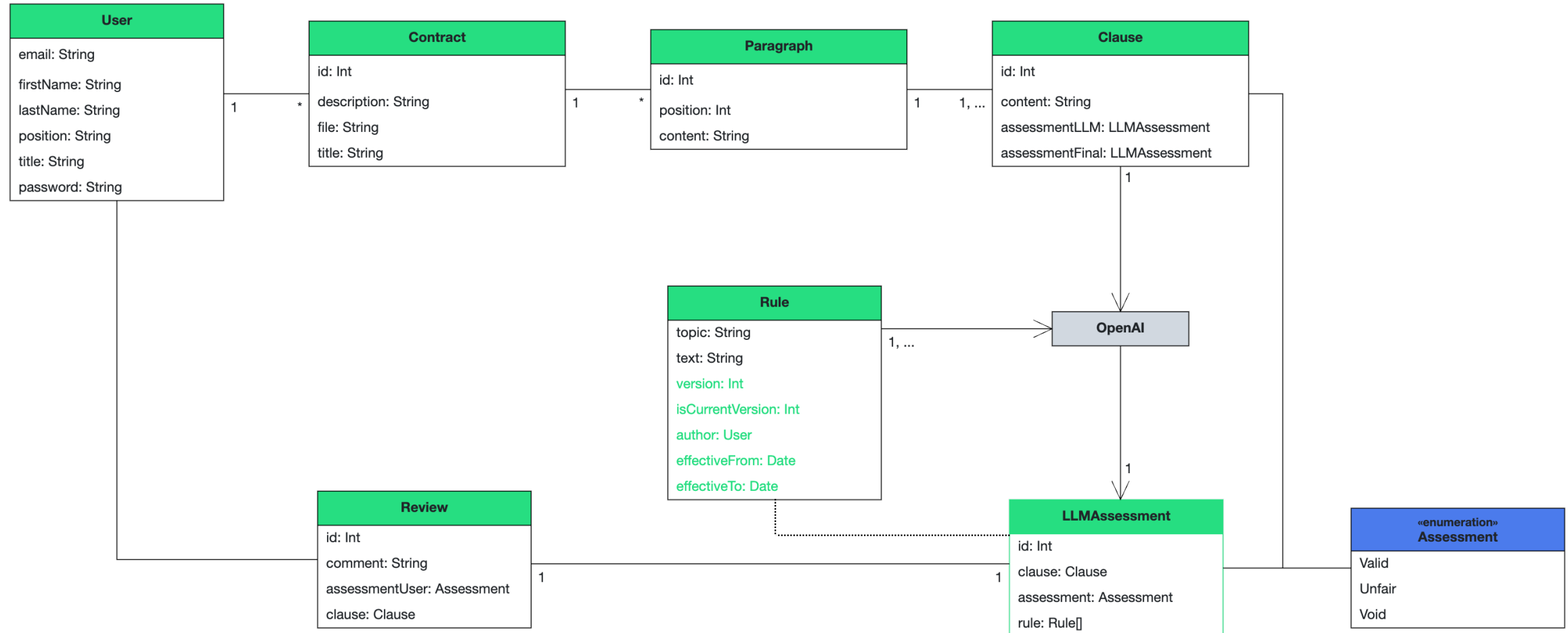
Background & Motivation

Current state of the system



Initial Findings

Potential changes of the system



Main Research Question:

Controlling the Evolution of a Legal Knowledge Base for an In-Context Learning AI System - Design and Implementation

RQ1

How can changes in examination guidelines and their impact on the evaluation of contract clauses in an in-context learning AI system be tracked and assessed?

RQ2

What is needed to create a user-friendly interface which allows to monitor this evolution without technical knowledge?

RQ3

How can technical aspects and the usability of the system be evaluated?

RQ1: How can changes in examination guidelines and their impact on the evaluation of contract clauses in an in-context learning AI system be tracked and assessed?

Current Prompt Structure

- The AI acts as a German lawyer specializing in employment contract law.
- The AI follows the provided rules (Prüfanleitungen) to evaluate the clause.
- The rules are assumed to be correct as they are provided by professional lawyers.
 - If a rule is violated → the clause is „**unzulässig**“.
- The response ends with one of three labels: "**unzulässig**," "**unfair**," or "**unbedenklich**."

Potential Solutions

1. Add final line with „**As last sentence return the applied rules.**“
2. Combine Chain-of-Thought Prompting with structured output
3. RAG system that filters applicable rules

[3, 9]

Main Research Question:

Controlling the Evolution of a Legal Knowledge Base for an In-Context Learning AI System - Design and Implementation

RQ1

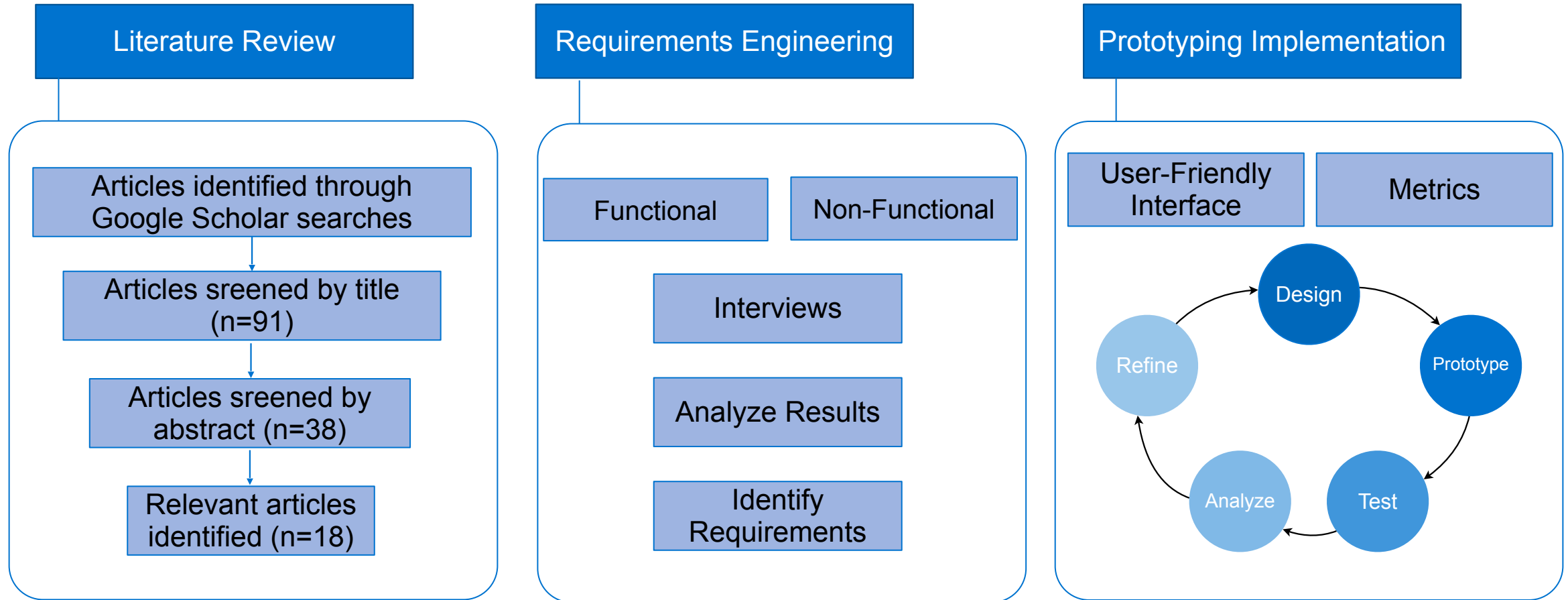
How can changes in examination guidelines and their impact on the evaluation of contract clauses in an in-context learning AI system be tracked and assessed?

RQ2

What is needed to create a user-friendly interface which allows to monitor this evolution without technical knowledge?

RQ3

How can technical aspects and the usability of the system be evaluated?

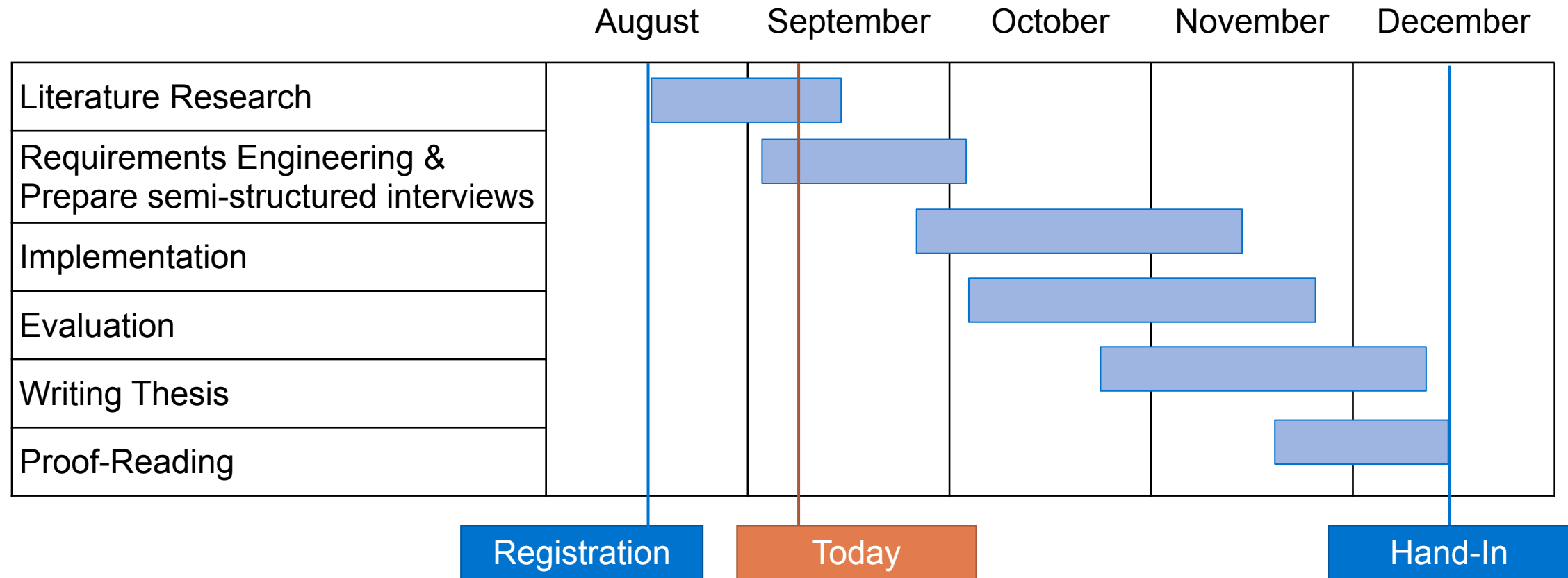


* until 06.09.2024

[5] Grimaldi (2021)

[4] Iterative design

Timeline





Prof. Dr.

Florian Matthes

Technical University of Munich (TUM)
TUM School of CIT
Department of Computer Science (CS)
Chair of Software Engineering for Business
Information Systems (sebis)

Boltzmannstraße 3
85748 Garching bei München

+49.89.289.

17132
www.matthes.in.tum.de
matthes@in.tum.de



- [1] M. Mosbach, T. Pimentel, S. Ravfogel, D. Klakow, and Y. Elazar, "Few-shot Fine-tuning vs. In-context Learning: A Fair Comparison and Evaluation," arXiv, May 26, 2023. <https://arxiv.org/abs/2305.16938>.
- [2] J. Wei et al., "Chain-of-Thought Prompting Elicits Reasoning in Large Language Models," arXiv, Jan. 28, 2022. <https://arxiv.org/abs/2201.11903>
- [3] OpenAI, "Structured Outputs Examples," OpenAI Documentation, [Online]. Available: <https://platform.openai.com/docs/guides/structured-outputs/examples?context=ex2>. [Accessed: 18-Sep-2024].
- [4] I. Xie and K. K. Matusiak, "Interface design and evaluation," in *Discover Digital Libraries*, Academic Press, 2016, pp. 255-297.
- [5] E. Grimaldi, "Requirements Engineering Methods and Techniques," Medium, Oct. 12, 2021. <https://medium.com/@edgargrimaldi/requirements-engineering-methods-and-techniques-802a20efaac7>
- [6] G. Tonye, "Machine Learning Confidence Scores: All You Need to Know as a Conversation Designer," **Voice Tech Global** on Medium, Aug. 25, 2021. <https://medium.com/voice-tech-global/machine-learning-confidence-scores-all-you-need-to-know-as-a-conversation-designer-8babd39caae7>
- [7] K. Tian, E. Mitchell, A. Zhou, A. Sharma, R. Rafailov, H. Yao, C. Finn, and C. D. Manning, "Just Ask for Calibration: Strategies for Eliciting Calibrated Confidence Scores from Language Models Fine-Tuned with Human Feedback," arXiv:2305.14975v2 [cs.CL], Oct. 2023.
- [8] D. Dalalah and O. M. A. Dalalah, "The false positives and false negatives of generative AI detection tools in education and academic research: The case of ChatGPT," **Computers and Education: Artificial Intelligence**, vol. 4, May 26, 2023. <https://www.sciencedirect.com/science/article/pii/S1472811723000605>
- [9] T. DD, "Enforcing Structured Output in Language Models," *Medium*, Aug. 21, 2024. <https://medium.com/@tamar-dd/enforcing-structured-output-in-language-models-dedee56ee729>
- [10] C. Marco et al., "Applicable law to a contract: What happens when the law changes?," **Association of Corporate Counsel**, Nov. 23, 2015. <https://www.acc.com/resource-library/applicable-law-contract-what-happens-when-law-changes>



Prof. Dr.

Florian Matthes

Technical University of Munich (TUM)
TUM School of CIT
Department of Computer Science (CS)
Chair of Software Engineering for Business
Information Systems (sebis)

Boltzmannstraße 3
85748 Garching bei München

+49.89.289.

17132
www.matthes.in.tum.de
matthes@in.tum.de



Full Prompt

`system_prompt = (`

"Du bist ein deutscher Anwalt, welcher auf Wirtschaftsrecht spezialisiert ist und Klauseln aus Arbeitsverträgen hinsichtlich ihrer rechtlichen Zulässigkeit und Fairness bewertet. "

"Du denkst Schritt für Schritt und beachtest dabei die Regeln, welche dir zusammen mit der Klausel zur Verfügung gestellt werden. "

"Die Regeln stammen von professionellen Anwälten und sind somit immer als korrekt anzusehen. "

"Du bewertest eine Klausel nur dann als 'unzulässig', wenn sie gegen eine der gegebenen Regeln verstößt. "

"Wenn keine der gegebenen Regeln Anwendung findet, nimmst du eine eigene Einschätzung vor. Dabei bewertest du die Klausel aber nur entweder als 'unfair' oder 'unbedenklich'. Du bewertest die Klausel als 'unfair' wenn sie den Arbeitgeber oder den Arbeitnehmer unangemessen benachteiligt oder besonders ungewöhnliche Formulierungen enthält."

"Deine Antwort beendest du immer mit einer abschließenden Bewertung als: 'unzulässig', 'unfair' oder 'unbedenklich'."

`)`

`user_prompt = "Ist die folgende Klausel unzulässig, unfair oder unbedenklich?`

`Klausel: ,###clause###'`

`Beachte bitte folgende Regeln: ###rules###"`

Method	Template
Label prob.	Provide your best guess for the following question. Give ONLY the guess, no other words or explanation.\n\nFor example:\n\nGuess: <most likely guess, as short as possible; not a complete sentence, just the guess!>\n\nThe question is:\${THE_QUESTION}
'Is True' prob.	Question: \${QUESTION}\nProposed Answer: \${ANSWER}\nIs the proposed answer:\n\t(A) True or\n\t(B) False?\n The proposed answer is:
Verb. 1S top-1	Provide your best guess and the probability that it is correct (0.0 to 1.0) for the following question. Give ONLY the guess and probability, no other words or explanation. For example:\n\nGuess: <most likely guess, as short as possible; not a complete sentence, just the guess!>\n Probability: <the probability between 0.0 and 1.0 that your guess is correct, without any extra commentary whatsoever; just the probability!>\n\nThe question is: \${THE_QUESTION}
Verb. 1S top- k	Provide your k best guesses and the probability that each is correct (0.0 to 1.0) for the following question. Give ONLY the guesses and probabilities, no other words or explanation. For example:\n\nG1: <first most likely guess, as short as possible; not a complete sentence, just the guess!>\n\nP1: <the probability between 0.0 and 1.0 that G1 is correct, without any extra commentary whatsoever; just the probability!> ... G k : < k -th most likely guess, as short as possible; not a complete sentence, just the guess!>\n\nP k : <the probability between 0.0 and 1.0 that G k is correct, without any extra commentary whatsoever; just the probability!> \n\nThe question is: \${THE_QUESTION}