

XHAILe: Explainable Hybrid AI for Computational Law and Accurate Legal Chatbots

Thomas Hildebrandt^{1,*}, Hugo A. López^{2,*}, Konstantinos Varvoutas², Henrik Palmer Olsen³, Marieke Anne Heyl³, Manex Aguirrezabal Zabaleta⁴, Bolette S. Pedersen⁴, Daniel Hershcovich¹, Ilias Chalkidis¹, Ali Mohammed Ali Al-Laith⁴, Amogh Raina¹, Tilman Zuckmantel¹, Morten Marquard⁵, Søren Debois⁵, Håkon Norman⁵, Leif Storgaard⁶, Ghita Thiesen⁷, Jim Foli⁷ and Jakob Eiby⁷

¹Department of Computer Science, Copenhagen University, Denmark

²Department of Applied Mathematics and Computer Science, Technical University of Denmark

³Faculty of Law, Copenhagen University, Denmark

⁴Department of Nordic Studies and Linguistics, Copenhagen University, Denmark

⁵DCR Solutions, Denmark

⁶TietoEvy, Denmark

⁷Kommuneforlaget, Local Government Denmark (KL), Denmark

Abstract

The XHAILe Grand Solutions project brings together leading Danish researchers in law, language technology, AI and computational law and providers of digital public administration and legal tech software, to develop accurate and explainable Hybrid-AI solutions that combine efficient AI-based translation of law to an open computational law rule format and a conversational agent using a rule engine via an open API. Here come trustworthy, new legal benchmarks and guidelines for the use of the technology. The results will be evaluated and demonstrated on use cases from local and central government.

Keywords

Hybrid AI, Computational Law, Legal Chatbots, DCR Graphs, NLP, LLM, eGovernment

1. Introduction

Our society has become highly dependent on the use of information systems to manage the interactions between citizens, companies and the public sector. According to [1], Danish companies in 2023 used 40 million hours a year alone for documentation and reporting to public authorities required by law, corresponding to DKK 23.5 billion every year [1]. At the same time, a large number of new regulations and changes to regulations is every year being imposed both at EU level and member state levels, on top of an already complex body of regulations. For instance, 3 new legal requirements were imposed on companies in the EU every week in the period from 2017-2022 [1] and the Danish legislation on social services consisted in 2020 of 2,213 clauses, which was changed 4.686 times from primo 2007 to mid 2019 [2].

The recent advances in the development of large language models (LLMs) and the rollout of commercial LLM-based chatbot assistants have raised expectations for the potential impact of such technologies across industries on interaction between governments, citizens and companies, not least with respect to the handling of legal matters. While in some use cases, end-to-end LLM-based solutions can be quite effective, if the user is aware of and ready to cope with the potential errors in the answers, the

Research Project Exhibition, CAISE 2026

*Corresponding authors and presenters.

✉ hilde@di.ku.dk (T. Hildebrandt); hulo@dtu.dk (H. A. López); konva@dtu.dk (K. Varvoutas); henrik@jur.ku.dk (H. P. Olsen); marieke@jur.ku.dk (M. A. Heyl); manex.aguirrezabal@hum.ku.dk (M. A. Zabaleta); bspedersen@hum.ku.dk (B. S. Pedersen); dh@di.ku.dk (D. Hershcovich); ilias.chalkidis@di.ku.dk (I. Chalkidis); alal@di.ku.dk (A. M. A. Al-Laith); amogh.raina@di.ku.dk (A. Raina); @di.ku.dk (T. Zuckmantel); morten@dcrsolutions.net (M. Marquard); debois@dcrsolutions.net (S. Debois); normann@DCRSolutions.net (H. Norman); leif.storgaard@tietoevy.com (L. Storgaard); GHT@kl.dk (G. Thiesen); JFO@Kommuneforlaget.dk (J. Foli); JAEI@kl.dk (J. Eiby)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

situation is quite different in high-stakes scenarios, such as those commonly encountered in processes that involve the interpretation and execution of law.

In the XHAILe project, we aim to address these challenges and develop trustworthy, legally compliant conversational legal assistants and computational law. The central idea is to treat LLMs as a bounded conversational interface to a rule-based *controller* that executes case-management logic and computational models of the law. As core technology for representing computational law we build on the formal declarative language of Dynamic Condition Response (DCR) graphs [3, 4].

Concretely, the objectives are the following:

1. Open-source prototype of rule engine for computational law format, which is consistent with international standards (LegalRuleML) and the DCR Solutions platform, able to express the computational rules in the selected use cases, and accessible by an open API
2. Open-source prototype of algorithm that can automatically translate a majority of the computational rules in each use case and a GUI that allows legal experts (who has received a few days of training) to complete the remaining part of the translation in less than an hour.
3. Open-source prototype of Hybrid-AI chatbot for conversational natural language interaction with e-government solutions that for each use case can apply the rule engine to provide accurate answers attributed to the relevant law understandable by non-legal experts.
4. Legal practice guidelines and trustworthy benchmarks for computational law and conversational interfaces for e-government processes.
5. Prototypes, practices and guidelines demonstrated at Society Readiness Level (SRL) and Technology Readiness Level (TRL) 7 on the use-cases in collaboration with central and local government, covering national laws in Danish and Norwegian as well as regulations at EU level with national implementations.

The project is planned for 3.5 years, running from April 2025 to October 2028.

2. Example: Application for Technical Aids and Consumer Durables

As an illustrative example, we use the right to get support for technical aids according to the Danish Consolidation Act on Social Services.¹ The rights are described in four sections: §112, §112a, §113, §113a and §113b. The law is intended to be managed digitally, indeed, the law stipulates in §112a, (1-3), *applications not submitted by means of digital self-service shall be rejected by the municipal council, unless special circumstances apply or if, based on an overall financial assessment, it is clearly an advantage for the municipal council to receive the application in a non-digital manner*. An example of a special circumstance can be, that the citizen is exempted from using digital solutions. Local Government Denmark maintains carefully developed official application forms as pdf documents, that both serve as requirements for the development of digital services and in special cases can be obtained from the municipality (or printed) and filled out manually. Besides the fields to be filled with information by the applicant, the forms include the relevant sections of the law. The fields relates to the conditions stipulated in the law. For instance, the law states in the beginning of §112 that: *The municipal council shall grant support for technical aids to persons who have a permanent physical or mental impairment*. This condition is represented by three tick-boxes where the citizen must answer *yes*, *no* or *don't know* to whether the citizen suffers from permanent physical or mental impairment, and a text box where the kind of impairment can be elaborated. The law also specifies in §112 (i-iii) three sub conditions, but instead of representing these conditions directly in the form, the form instead contains text boxes, where the citizen can describe in natural language which challenges in the daily life that are the reason for the application.

The use of the forms as requirement specifications for the digital self-service solutions means, that (at least) two manual translations are done during the development of the digital self-service solutions, as depicted in Fig. 1. Each of these translations involve interpretations of the law, notably on what data

¹Unofficial English translation available here: www.english.sm.dk/media/14900/consolidation-act-on-social-services.pdf

is required to be provided by the citizen, under which conditions, and what information the citizen must be given (and consent to) as part of the application process.

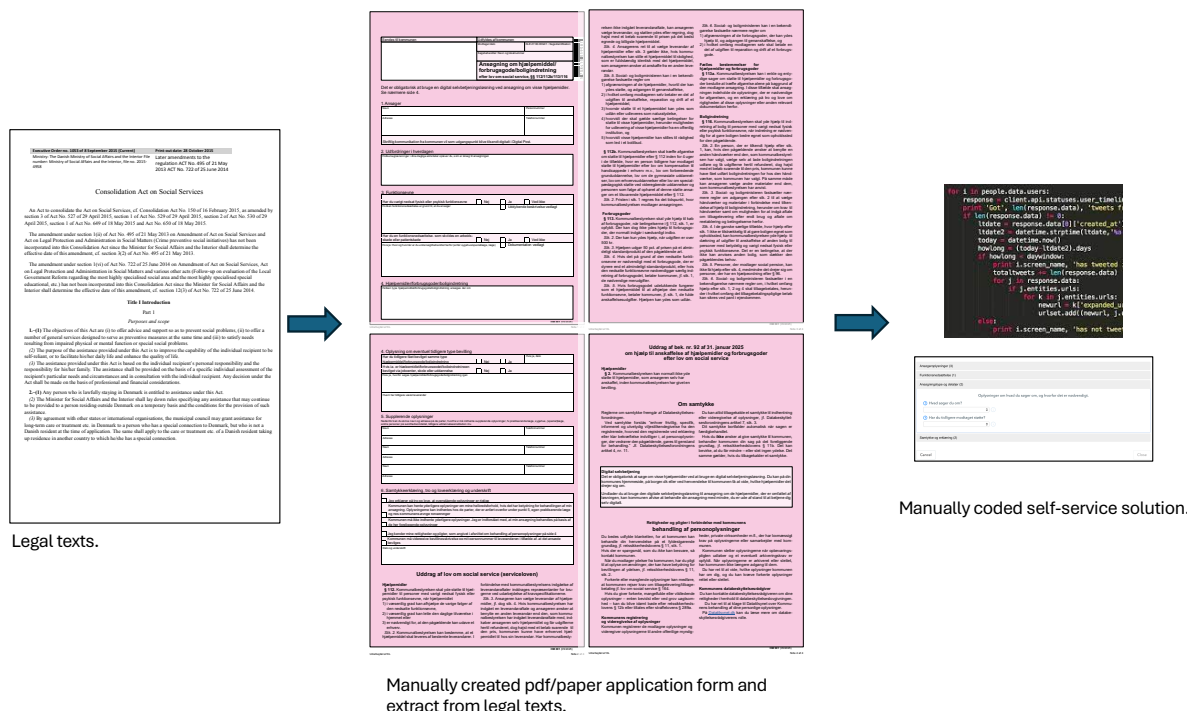


Figure 1: Laws are today manually translated to official pdf forms that are then manually coded as self-service solutions. The text (in Danish) and the code (not the real code of a self-service solution) are for illustration purposes only, and are not meant to be readable.

3. Hybrid AI for Computational Law and Accurate Legal Chatbots

In XHAILe we investigate the use of a *hybrid AI* architecture, where sub-symbolic models (e.g. language models) are used as translators of inputs from users or knowledge bases and symbolic models (e.g. rule based models) are used as controllers of the interaction and to infer the answer.

As foundation for the symbolic rules, we build on the work of *Dynamic Condition Response (DCR) Graphs*, an executable, declarative, temporal logic process modeling language [3]. Unlike imperative process modelling notations (like BPMN [5]) that rigidly describe the sequencing of possible steps as flow charts, DCR graphs model the rules. This aligns naturally with the nature of law, which typically prescribes obligations and conditions, rather than scripting it. The modelling language of DCR graphs was introduced in [3]. Since then, the language has been extended to be able to model processes with time [6], data and decision modelling [4], introducing data-dependent rules and embedded computations, and also integrating the Decision Model and Notation (DMN) standard [7]. The task of modelling is also supported by commercial cloud-based solutions, including tools for simulation of processes and for assisting domain and legal experts in the mapping of text based process models, guidelines, and regulations into DCR graphs [8]. DCR graphs have been successfully deployed in Danish public-sector case-management systems using a cloud-based process-engine-as-a-service, that is integrated into commercial platforms [9].

The DCR technology and tools, however, still have some limitations that need to be addressed as part of the research project, which is elaborated on below.

3.1. Scalable DCR Graphs for Law

Grouping of activities that have rules in common can help reducing the complexity [10]. Also, initial work on separating DCR graphs in sub parts inspired by the modelling of law was initiated in [11], introducing networks of DCR graphs and linking of activities between sub parts. We propose extending this direction of work to investigate structure specifically relevant for the modelling of law. Concretely, we aim to explore the structure of legal documents and standardised formats, in particular the Akoma Ntoso for European Union (AKN4EU) format,² which is chosen to be the basis for exchange of legal documents in the EU. We will also explore the concepts for reuse from object oriented modelling in the context of symbolic models of law. Initial work on object-centric DCR graphs was initiated in [12, 13]. We propose to extend this line of work by investigating suitable definitions of inheritance between classes of DCR graphs that still respect the principles of encapsulation [14] and investigate their usefulness in the modelling of law.

3.2. Translating Law to DCR Graphs

The current state of the art of using NLP techniques for transforming texts to DCR Graphs [8] completely ignores data. To support form filling (and possibly replacing it with conversations), the DCR graphs should contain nodes corresponding to the input fields and rules representing which information is relevant or required based on the data provided in the input fields. In our working example, we already saw examples of data values, such as the type of impairment that can be either physical or mental. Another example found in §113 (2), which gives a data-dependent rule: *Assistance shall only be provided where the cost exceeds DKK 500* which is based on a data value representing the *cost* of the consumer durables.

In the area of software engineering, there exists a number of relevant works for translating software requirements to object class and data models, spanning from work using classical NLP techniques (e.g. [15, 16]) to more recent work based on LLMs (e.g. [17, 18]). Similarly, baseline works on the extraction of imperative process models from texts consider the activity data as part of their annotation schemes [19]. We propose to investigate these strands of work as a starting point for including data and data dependencies in the mapping of legal texts, including both regulations and guidelines, to DCR graphs. This will also go hand-in-hand with the exploration of object oriented concepts for tackling the scalability challenge for DCR graph models.

To support the subsequent case management, the DCR graphs should also contain information about the ordering and timing of actions to be taken, which again can depend on data. The existing work in [8] captures some ordering constraints, but ignores data and timing altogether. Also, as mentioned above, the translations were targeting descriptions of business processes and not legal texts. To this end, we aim to explore existing work on translations of descriptions of processes with both data and timing constraints, in particular in the legal domain.

There has been extensive work on building knowledge graphs from natural language [20, 21]. These approaches, though, extract encyclopedic knowledge, such as capital/country relations, river locations, birthplaces of famous people, and so on. In our case, the type of knowledge and relations that we are interested in does not have that same shape. We are interested in events, actors, and their relations, with respect to the events described in laws and legal guidelines. This has been approached in meaning representation and semantic parsing work, such as [22, 23], which develops models to extract events, entities, and relations between them—including, specifically, from legal text [24, 25, 26, 27]. We take inspiration from this work, treating legal process extraction as a semantic parsing task.

3.3. An Accurate Conversational Legal Chatbot

We propose a *flexible, controllable, and legally compliant* Conversational Hybrid AI (CHAI) assistant that as shown in Fig. 2 treats large (or if sufficient, small) language models strictly as a *bounded*

²<https://op.europa.eu/en/web/eu-vocabularies/akn4eu>

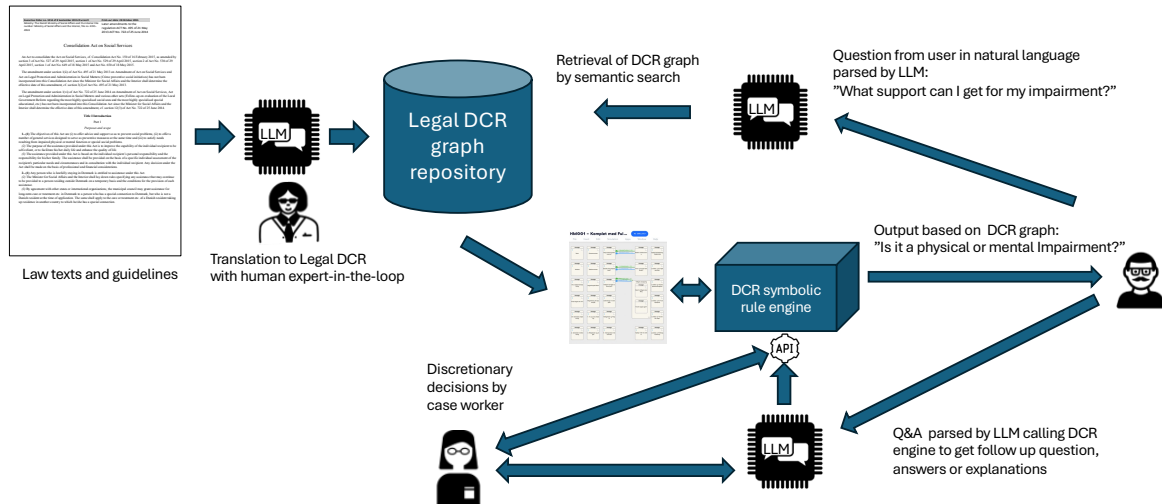


Figure 2: Conversational Hybrid AI (CHAI) Architecture: LLMs are used for parsing input and searching for Legal DCR Graphs, which are then used to provide answers and explanations using the API of the Extended DCR Rule Engine.

interface to a rule-based *controller* (similar to the proposal in [28]) executing legal decision and dynamic process logic encoded in (an extension of) the declarative DCR graphs process notation, extracted from legal regulations and guidelines by using classic NLP techniques and LLMs, with legal experts as human-in-the-loop.

The assistant is initially scoped to an *Interpreter for submission and management of applications from citizens to governmental agencies*: It elicits and normalizes user inputs, maps them to the input fields and actions in the logical process model, and presents controller-backed explanations, while *all* execution and decisions remain with the controller. This separation maintains public-authority accountability: the controller guarantees sound execution, provenance, and auditability; the LLM improves accessibility (intent understanding and language adaptation) but does not determine outcomes. A dialogue loop can be realized in which the LLM converts free-form utterances into schema-constrained structures aligned with the controller’s inputs, queries the controller to test preconditions and obligations, and renders the controller’s trace and explanation into citizen-facing feedback. Explanations are fact–rule disaggregated: the interface states what the user supplied, what the controller verified, which conditions are satisfied or unmet, and which next steps follow from the model.

We envisage three iterations. In the near term, we formalize the dialogue tool grammar (schema-validated calls with explicit uncertainty fields), extend the controller interface with functions for suggesting required fields and next steps, and standardize fact–rule faithful explanations over controller traces. We also introduce confidence-aware verification, limiting confirmations and paraphrase-backs to cases flagged by uncertainty or risk. In the medium term, we add process routing via an intent router trained on process metadata and entry conditions, enforced by hard guardrails that prevent calls outside an activated process. Language and persona policies are introduced as presentation constraints that never alter execution semantics and are logged alongside traces; escalation pathways provide succinct, trace-linked summaries for caseworkers.

3.4. Legal Implications

Using DCR graphs for legal decision-making can only be implemented if the graphs and their use is fully compliant with the law as a whole. Building a graph to support and partially automate steps for a specific provision in law (here articles 112-113 in the social services act) need to take into account the entire legal framework relevant to ensure rightful administration. This entails among other things that

the provisions in the Danish Public Administration Act³ are taken into consideration where relevant (this entails for example a hearing of the citizen before a final decision is made); that relevant case law, either from Danish Courts or the National Social Appeals Board⁴ are taken into account (The general principle of equality before the law, entails that previous decisions influence future decisions in similar cases); and that instructions internal to the administration⁵ is followed (such instructions lays out interpretive guidelines which individual caseworkers are expected to follow in their decision-making practice).

Formalizing the administrative work process, by making each of the information retrieval and sub-decision steps visible in a graph has a potential to enhance both legal certainty and administrative efficiency, but it is important to distinguish between different use cases, primarily between citizen facing chatbots and caseworker facing chatbots. The main difference here is that citizen facing chatbots should be able to support any citizens seeking general information about how the legal rules is applied in administrative practice, while caseworker facing chatbots has as its purpose to assist caseworkers making binding decisions in a concrete case taking into consideration all the relevant facts of the case. In both cases, of course, the chatbots must provide lawful information and must be able to relate to inputted prompts in a trustworthy and robust way, but there are important differences that dictate separate set-ups for each of these two kinds of chatbots.

Chatbots for citizens can not be set up to provide suggestions for decisions as this could generate unfounded expectations in the citizen about the outcome of a case. Guidance to citizens therefore must take the form of what kind of information that is relevant to the decision and what kind of information that is irrelevant. It is also important to note that such information can never be entirely accurate, since what is relevant may depend on many different circumstances relevant to the specific case. We suggest that the best option here is to operationalise existing internal guidelines ("cirkulærer" and "vejledninger") in a guardrailed chatbot architecture built around the relevant rules.

Chatbots for caseworkers, on the other hand, could best be build in such a way that it provides suggestions for decisions relevant to each step in the decision-making process. By decomposing the legal rule into individual decision steps with the DCR tool, it becomes possible to use different solutions to provide suggestions. Depending on the kind of sub-decision, rule-based, expert-based or Rag-based solutions could be used to provide suggestions to the case worker. The DCR format will be specifically helpful here, as it allows the case-worker to move both forward and backwards between sub-decisions—a necessary mechanism for legal decision making that involve interpretation or discretion, since these are made on the basis of a holistic analysis of all elements in the case as a whole. To make a decision under §112 for example, it is necessary to assess the specific impairment with the specific layout of the person's home. While retrieving information about impairment and home may be seen as two separate steps in a case management flow, the final decision about whether to provide aid and and what kind of aid depends on the precise interaction between impairment and home in such a way that it is necessary to observe and analyse those two information steps under one single view.

4. Conclusion and Future Work

We have above described the roadmap to explainable and accurate legal chatbots to be followed in the XHAILe project. The roadmap builds on prior successful work on no-code digitalisation of law and eGovernment services using the DCR technologies, which is already widely used to digitalize governmental case management processes in Denmark. The described activities aim to advance the state of the art, by at the one hand formalizing the role of the LLMs in the use of eGovernment solutions as a bounded user interface to a DCR-based controller and at the other hand the use of language technology in the development and maintenance of eGovernment solutions for the translation of law texts and

³<https://www.retsinformation.dk/eli/lta/2014/433>

⁴The most important of the latter decisions are published here: <https://www.ast.dk/for-myndigheder/principmeddelelser>

⁵see for the §112 and §113 example we have used here, the instructions at: <https://www.retsinformation.dk/eli/retsinfo/2017/10324>

guidelines into DCR models. To safely execute the roadmap we have gathered an interdisciplinary team of legal scholars and computer scientists, spanning from LLMs and NLP to symbolic logic and technologies for case and business process management. Hereto comes partners from local and central governmental organisations providing access to relevant cases and giving feedback on the value and usefulness of the developed framework. This makes it possible to follow the roadmap using a Reflective System Development approach [29]. Reflective System Development approach is an action research method where researchers are actively engaged with practitioners when investigating opportunities and addressing challenges equally relevant for both research and practice. Reflective System Development has a distinct cyclic iterative nature formed by three main research activities: 1) understanding through interpretation, 2) supporting through design, and 3) improving through intervention. We have planned four cycles of 6 months, of which the first is currently in progress, taking the solution from TRL (Technology Readiness Level) and SRL (Society Readiness Level) 2 to 7, that is, the solution is tested in operational environment with relevant stakeholders.

Acknowledgments

This work is supported by Innovation Fund Denmark.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

- [1] C. Aastrup Seidelin, S. Tomori Oyebo, Papirarbejde er dyrt for danske virksomheder og deres medarbejdere – og regningen vokser, 2024. URL: <https://www.danskindustri.dk/arkiv/analyser/2024/9/papirarbejde-er-dyrt-for-danske-virksomheder-og-deres-medarbejdere--og-regningen-vokser/>.
- [2] B. Arent Eiriksson, I. Nordland, Regelkompleksitet på det socialretlige område, 2020. URL: <https://justitia-int.org/analyse-regelkompleksitet-paa-det-socialretlige-omraade/>.
- [3] T. T. Hildebrandt, R. R. Mukkamala, Declarative event-based workflow as distributed dynamic condition response graphs, arXiv preprint arXiv:1110.4161 (2011).
- [4] T. T. Hildebrandt, H. Normann, M. Marquard, S. Debois, T. Slaats, Decision modelling in timed dynamic condition response graphs with data, in: International Conference on Business Process Management, Springer, 2021, pp. 362–374.
- [5] Object Management Group, Business Process Model and Notation (BPMN), Version 2.0.2, Technical Report, Object Management Group, 2013. URL: <https://www.omg.org/spec/BPMN/2.0.2>, Formal Version: 2013-12-09.
- [6] T. T. Hildebrandt, R. R. Mukkamala, T. Slaats, F. Zanitti, Contracts for cross-organizational workflows as timed dynamic condition response graphs, J. Log. Algebr. Program. 82 (2013) 164–185.
- [7] O. M. Group, Decision Model and Notation V 1.3 (2020). URL: <https://www.omg.org/spec/DMN/1.3/>.
- [8] H. A. López, T. Hildebrandt, S. Debois, M. Marquard, The process highlighter: From texts to declarative processes and back, ceur workshop proceedings, 2018, pp. 66–70. Joint Proceedings of the BIR 2018 Short Papers, Workshops and Doctoral Consortium co-located with 17th International Conference Perspectives in Business Informatics Research (BIR 2018).
- [9] M. Marquard, T. Slaats, From law to dcr graphs: A case study in digitizing public administration, in: Business Process Management Workshops, Springer, 2017.
- [10] V. P. Cosma, A. K. F. Christfort, T. T. Hildebrandt, X. Lu, H. A. Reijers, T. Slaats, Improving simplicity by discovering nested groups in declarative models, in: G. Guizzardi, F. Santoro, H. Mouratidis, P. Soffer (Eds.), Advanced Information Systems Engineering, Springer Nature Switzerland, Cham, 2024, pp. 440–455.

- [11] H. A. López, S. Debois, T. Slaats, T. T. Hildebrandt, Business process compliance using reference models of law, in: H. Wehrheim, J. Cabot (Eds.), *Fundamental Approaches to Software Engineering*, Springer International Publishing, Cham, 2020, pp. 378–399.
- [12] A. K. Christfort, A. Rivkin, D. Fahland, T. T. Hildebrandt, T. Slaats, Discovery of object-centric declarative models, in: *2024 6th International Conference on Process Mining (ICPM)*, 2024, pp. 121–128. doi:10.1109/ICPM63005.2024.10680680.
- [13] T. Zuckmantel, Y. Zhou, B. Düdler, T. Hildebrandt, Daceo: Declarative asynchronous choreographies with general data-dependent event-ordering and objects, in: C. Di Giusto, A. Ravara (Eds.), *Coordination Models and Languages*, Springer Nature Switzerland, Cham, 2025, pp. 197–216.
- [14] A. Snyder, Encapsulation and inheritance in object-oriented programming languages, *SIGPLAN Not.* 21 (1986) 38–45. URL: <https://doi.org/10.1145/960112.28702>. doi:10.1145/960112.28702.
- [15] H. Harmain, R. Gaizauskas, Cm-builder: A natural language-based case tool for object-oriented analysis, *Automated Software Engineering* 10 (2003) 157–181. URL: <https://doi.org/10.1023/A:1022916028950>.
- [16] T. Yue, L. Briand, Y. Labiche, A systematic review of transformation approaches between user requirements and analysis models, *Requirements Engineering* 16 (2011) 75–99. URL: <https://doi.org/10.1007/s00766-010-0111-y>.
- [17] D. De Bari, G. Garaccione, R. Coppola, M. Torchiano, L. Ardito, Evaluating large language models in exercises of uml class diagram modeling, in: *Proceedings of the 18th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement, ESEM '24*, Association for Computing Machinery, New York, NY, USA, 2024, p. 393–399. URL: <https://doi.org/10.1145/3674805.3690741>. doi:10.1145/3674805.3690741.
- [18] K. Chen, Y. Yang, B. Chen, J. A. Hernández López, G. Mussbacher, D. Varró, Automated domain modeling with large language models: A comparative study, in: *2023 ACM/IEEE 26th International Conference on Model Driven Engineering Languages and Systems (MODELS)*, 2023, pp. 162–172. doi:10.1109/MODELS58315.2023.00037.
- [19] P. Bellan, H. van der Aa, M. Dragoni, C. Ghidini, S. P. Ponzetto, Pet: an annotated dataset for process extraction from natural language text tasks, in: *International Conference on Business Process Management*, Springer, 2022, pp. 315–321.
- [20] H. Ji, R. Grishman, H. T. Dang, K. Griffitt, J. Ellis, Overview of the TAC 2010 knowledge base population track, in: *Third text analysis conference (TAC 2010)*, volume 3, 2010, pp. 3–3.
- [21] J. Li, L. Qian, P. Liu, T. Liu, Construction of legal knowledge graph based on knowledge-enhanced large language models, *Information* 15 (2024). URL: <https://www.mdpi.com/2078-2489/15/11/666>. doi:10.3390/info15110666.
- [22] S. Oepen, O. Abend, J. Hajic, D. Hershcovich, M. Kuhlmann, T. O’Gorman, N. Xue, J. Chun, M. Straka, Z. Uresova, MRP 2019: Cross-framework meaning representation parsing, in: S. Oepen, O. Abend, J. Hajic, D. Hershcovich, M. Kuhlmann, T. O’Gorman, N. Xue (Eds.), *Proceedings of the Shared Task on Cross-Framework Meaning Representation Parsing at the 2019 Conference on Natural Language Learning*, Association for Computational Linguistics, Hong Kong, 2019, pp. 1–27. URL: <https://aclanthology.org/K19-2001>. doi:10.18653/v1/K19-2001.
- [23] S. Oepen, O. Abend, L. Abzianidze, J. Bos, J. Hajic, D. Hershcovich, B. Li, T. O’Gorman, N. Xue, D. Zeman, MRP 2020: The second shared task on cross-framework and cross-lingual meaning representation parsing, in: S. Oepen, O. Abend, L. Abzianidze, J. Bos, J. Hajič, D. Hershcovich, B. Li, T. O’Gorman, N. Xue, D. Zeman (Eds.), *Proceedings of the CoNLL 2020 Shared Task: Cross-Framework Meaning Representation Parsing*, Association for Computational Linguistics, Online, 2020, pp. 1–22. URL: <https://aclanthology.org/2020.conll-shared.1>. doi:10.18653/v1/2020.conll-shared.1.
- [24] S. T. Vu, M. Le Nguyen, K. Satoh, Abstract meaning representation for legal documents: an empirical research on a human-annotated dataset, *Artif. Intell. Law* 30 (2022) 221–243. URL: <https://doi.org/10.1007/s10506-021-09292-6>. doi:10.1007/s10506-021-09292-6.
- [25] A. Spangher, Z. Xue, T.-L. Wu, M. Hansen, J. May, Legaldiscourse: Interpreting when laws apply and to whom, in: *Proceedings of the 2024 Conference of the North American Chapter of the*

Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers), 2024, pp. 8528–8551.

- [26] N. Schrack, R. Cui, H. López, D. Hershcovich, Can AMR assist legal and logical reasoning?, in: Findings of the Association for Computational Linguistics: EMNLP 2022, Association for Computational Linguistics, Abu Dhabi, United Arab Emirates, 2022, pp. 1555–1568. URL: <https://aclanthology.org/2022.findings-emnlp.112>.
- [27] S. Vijay, D. Hershcovich, Can Abstract Meaning Representation facilitate fair legal judgement predictions?, in: S. Tafreshi, A. Akula, J. Sedoc, A. Drozd, A. Rogers, A. Rumshisky (Eds.), Proceedings of the Fifth Workshop on Insights from Negative Results in NLP, Association for Computational Linguistics, Mexico City, Mexico, 2024, pp. 101–109. URL: <https://aclanthology.org/2024.insights-1.13>.
- [28] P. Belcak, G. Heinrich, S. Diao, Y. Fu, X. Dong, S. Muralidharan, Y. C. Lin, P. Molchanov, Small language models are the future of agentic ai, 2025. URL: <https://arxiv.org/abs/2506.02153>. arXiv:2506.02153.
- [29] L. Mathiassen, Reflective systems development, Scandinavian Journal of Information Systems 10 (1998). URL: <https://aisel.aisnet.org/sjis/vol10/iss1/12>. doi:10.17705/3sjis.01012.