Boosting productivity using GenAI & NLP

Romy van Drie | TNO

TNO innovation for life

Content

- Evolution of NLP

- NLP work in Norm Engineering

  - Extracting components of FLINT frames

  - Interacting with FLINT frames

Evolution of NLP

Basis of NLP

Distributional hypothesis:

Words that occur in the same contexts tend to have similar meanings (Harris, 1954)

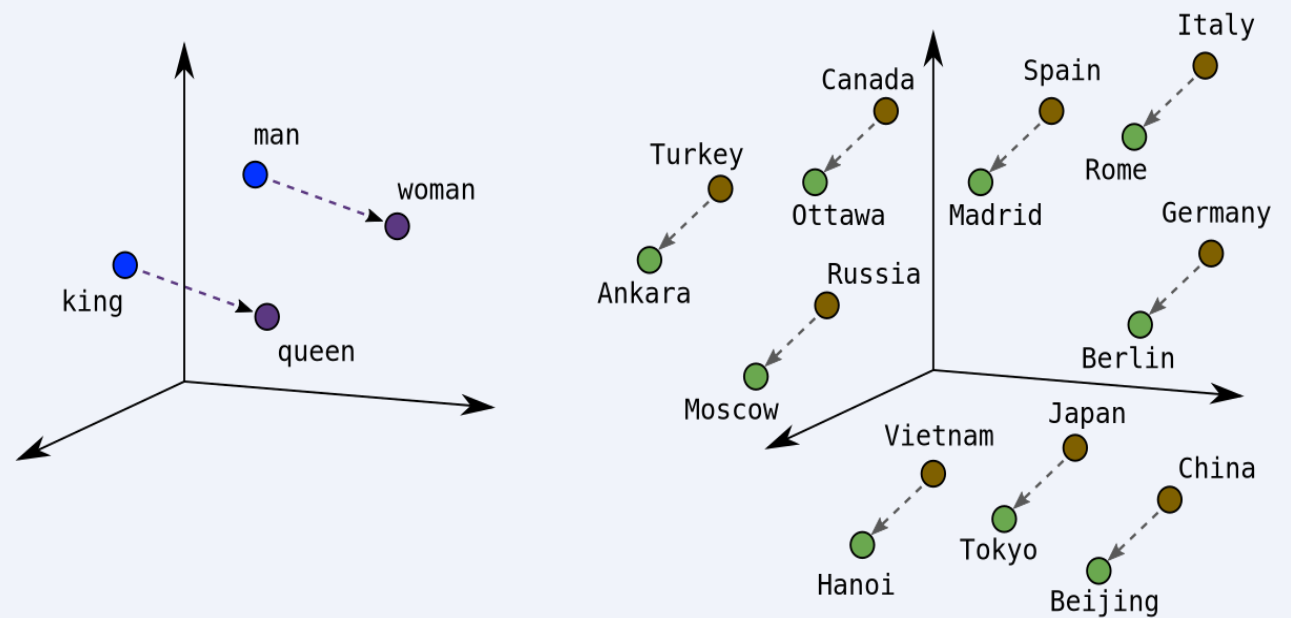TNO innovation for life

Evolution of NLP

Basis of NLP

Distributional hypothesis:

Words that occur in the same contexts tend to have similar meanings (Harris, 1954)

2013: word2vec (Mikolov et al.)

Provided high quality word embeddings

Evolution of NLP

Basis of NLP

Distributional hypothesis:

Words that occur in the same contexts tend to have similar meanings (Harris, 1954)

2013: word2vec (Mikolov et al.)

Provided high quality word embeddings

2018: BERT (Devlin et al.)

Provided a pre-trained model that could be finetuned for a wide variety of downstream NLP tasks (e.g. sentiment analysis)

Evolution of NLP

Basis of NLP

Distributional hypothesis:

Words that occur in the same contexts tend to have similar meanings (Harris, 1954)

2013: word2vec (Mikolov et al.)

Provided high quality word embeddings

2018: BERT (Devlin et al.)

Provided a pre-trained model that could be finetuned for a wide variety of downstream NLP tasks (e.g. sentiment analysis)

2020: GPT-3 (Brown et al.)

Provided generative capabilities

Can perform tasks with minimal or no task-specific training

Evolution of NLP

Basis of NLP

Distributional hypothesis:

Words that occur in the same contexts tend to have similar meanings (Harris, 1954)

2013: word2vec (Mikolov et al.)

Provided high quality word embeddings

2018: BERT (Devlin et al.)

Provided a pre-trained model that could be finetuned for a wide variety of downstream NLP tasks (e.g. sentiment analysis)

2020: GPT-3 (Brown et al.)

Provided generative capabilities

Can perform tasks with minimal or no task-specific training

2022: ChatGPT

Made GenAI accessible to the general public

TNO innovation for life

Evolution of NLP

Basis of NLP

2013: word2vec
(Mikolov et al.)

2018: BERT
(Devlin et al.)

2020: GPT-3
(Brown et al.)

2022: ChatGPT

Distributional hypothesis:

Words that occur in the same contexts tend to have similar meanings (Harris, 1954)

Provided high quality word embeddings

Provided a pre-trained model that could be finetuned for a wide variety of downstream NLP tasks (e.g. sentiment analysis)

Provided generative capabilities

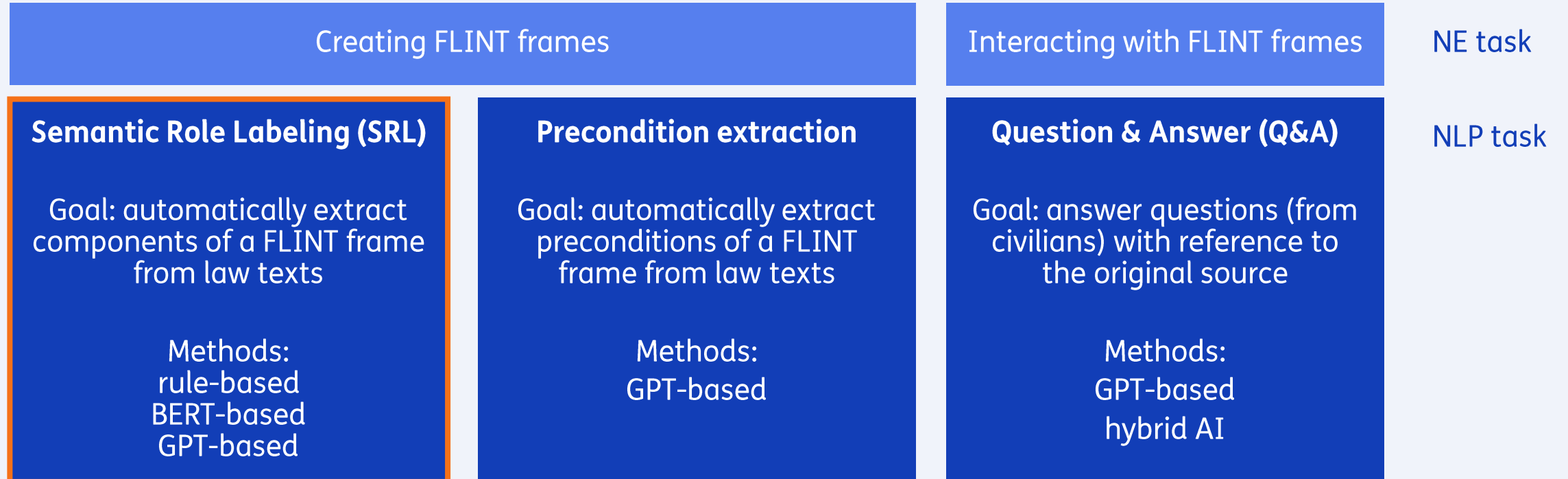Can perform tasks with minimal or no task-specific training

Made GenAI accessible to the general public

What does this mean for Norm Engineering?
Tasks that were previously 'impossible' now become possible!

TNO innovation for life

How can we use NLP for Norm Engineering?

Creating FLINT frames		Interacting with FLINT frames	NE task

| **Semantic Role Labeling (SRL)**

Goal: automatically extract components of a FLINT frame from law texts

Methods:
rule-based
BERT-based
GPT-based | **Precondition extraction**

Goal: automatically extract preconditions of a FLINT frame from law texts

Methods:
GPT-based | **Question & Answer (Q&A)**

Goal: answer questions (from civilians) with reference to the original source

Methods:
GPT-based
hybrid AI | NLP task |

TNO innovation for life

How can we use Semantic Role Labeling for NE?

Act frame

The initiator of some action, capable of acting with volition

The entity undergoing the effect of the action

The entity for whom the action was performed

Component	Example
Action	process
Actor	processor
Object	personal data
Recipient	data subject
Precondition	–
Postcondition	–
Source text	Personal data shall be processed lawfully, fairly and in a transparent manner in relation to the data subject
Source	Art. 5 (1) GDPR

TNO innovation for life

Method and results

Method

- Rule-based (Bakker et al., 2022a)

- BERT-based (van Drie et al., 2023; Bakker et al. 2022b)

- GPT-based (paper under review)

Results

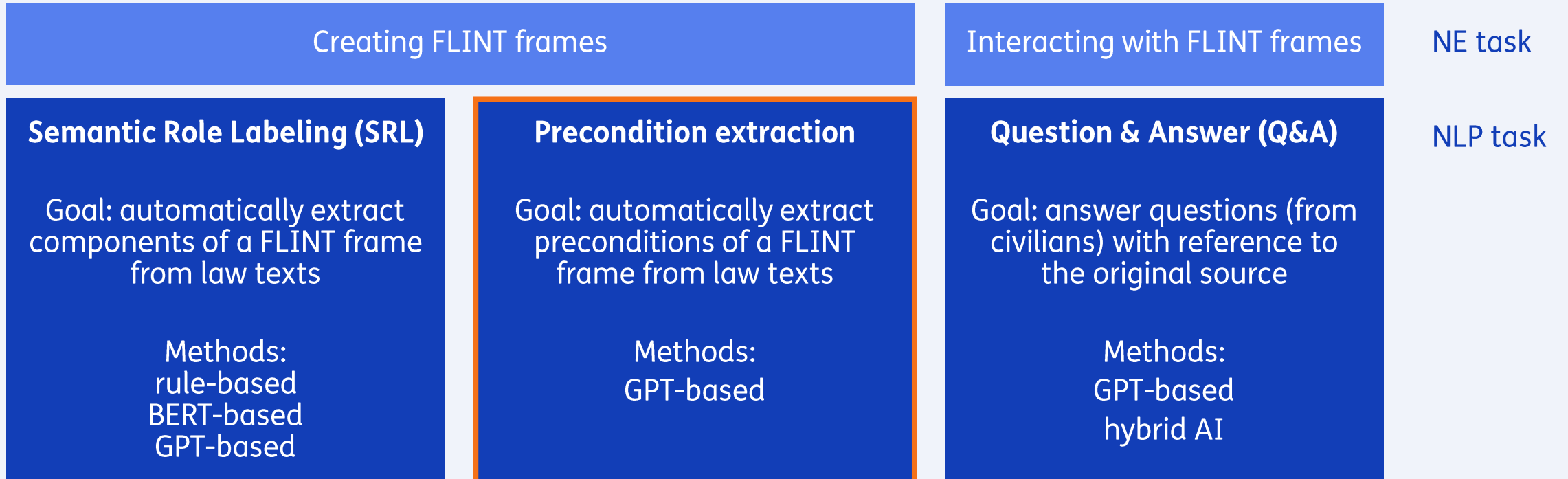|	Rule	BERT(je)	GPT-3.5 Turbo
English	0.528	**0.879**	0.627
Dutch	0.587	**0.842**	0.644

Fig. 1: accuracy

(out of all predictions, how many were correct?)

Conclusion

- BERT-based method still outperformed GPT-based method at time of testing

- However, BERT-based method needs considerable training data

  - English dataset: ~1500 labeled sentences

  - Dutch dataset: ~4500 labeled sentences

- It is likely that the GPT-based method will continue to improve as the GPT models improve

TNO innovation for life

How can we use Precondition Extraction for NE?

Act frame

Specific requirements, criteria, or circumstances that must be fulfilled before a specific action can occur

Component	Example
Action	terminate
Actor	the court
Object	guardianship
Recipient	-
Precondition	the guardian abuses authority or lacks required consent
Postcondition	-
Source text (simplified)	The court can terminate guardianship if the guardian abuses authority or lacks required consent
Source	Book 1 Dutch Civil Code, Article 327

TNO innovation for life

Method and results
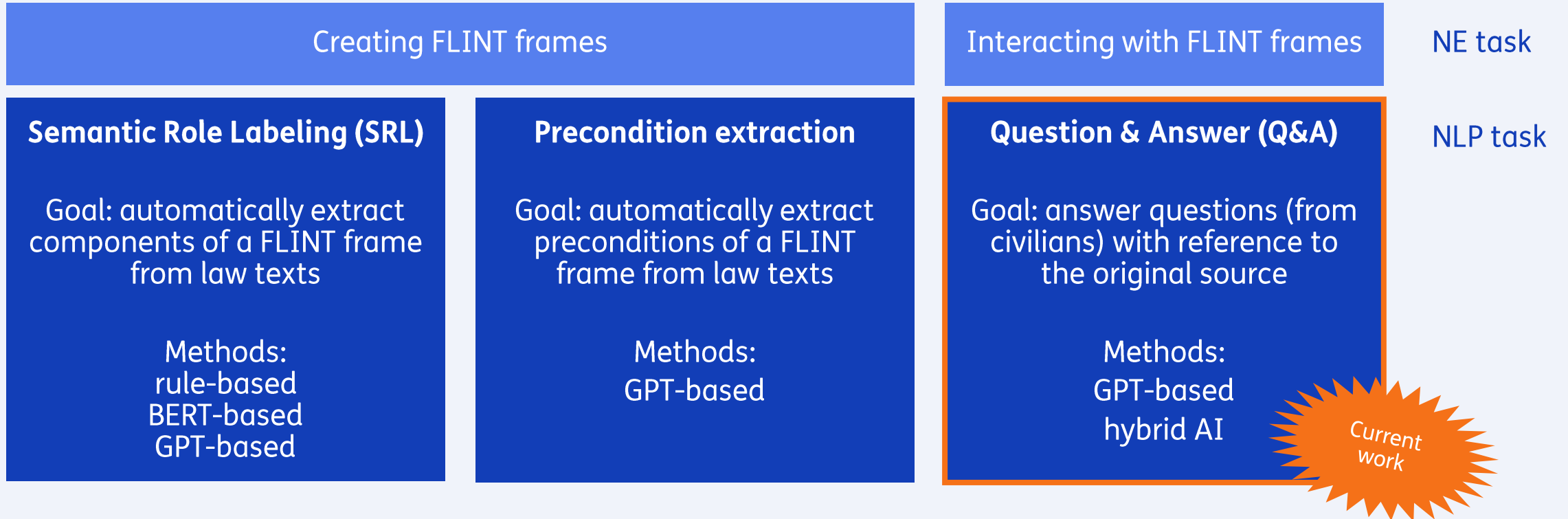
(Redelaar, van Drie, Verberne & de Boer, 2024)

Method

- Creation of dataset (102 legal QA pairs with source)

  - Question: "When can the Dutch court terminate the guardianship of a natural person?"

  - Answer: "The court can terminate guardianship if the guardian abuses authority or lacks required consent"

  - Source: "Book 1 Dutch Civil Code, Article 327"

- Implementation of LLM pipeline (RAG)

- Implementation of evaluation metrics (e.g. fluency, correctness, citation quality)

Conclusion

- It is possible to generate fluent and correct answers, with moderately reliable citation (83%)

- Extra work is needed to connect the QA-system to give suggestions for FLINT frames

TNO innovation for life

How can we use NLP for Norm Engineering?

Creating FLINT frames		Interacting with FLINT frames	NE task

Semantic Role Labeling (SRL)	**Precondition extraction**	**Question & Answer (Q&A)**	NLP task
Goal: automatically extract components of a FLINT frame from law texts	Goal: automatically extract preconditions of a FLINT frame from law texts	Goal: answer questions (from civilians) with reference to the original source	
Methods: rule-based BERT-based GPT-based	Methods: GPT-based	Methods: GPT-based hybrid AI	

Current work

TNO innovation for life

How can people interact with FLINT?

- Representation in FLINT can quickly become quite large, especially in complex situations

- Can we think of additional ways to interact with FLINT to improve user friendliness?
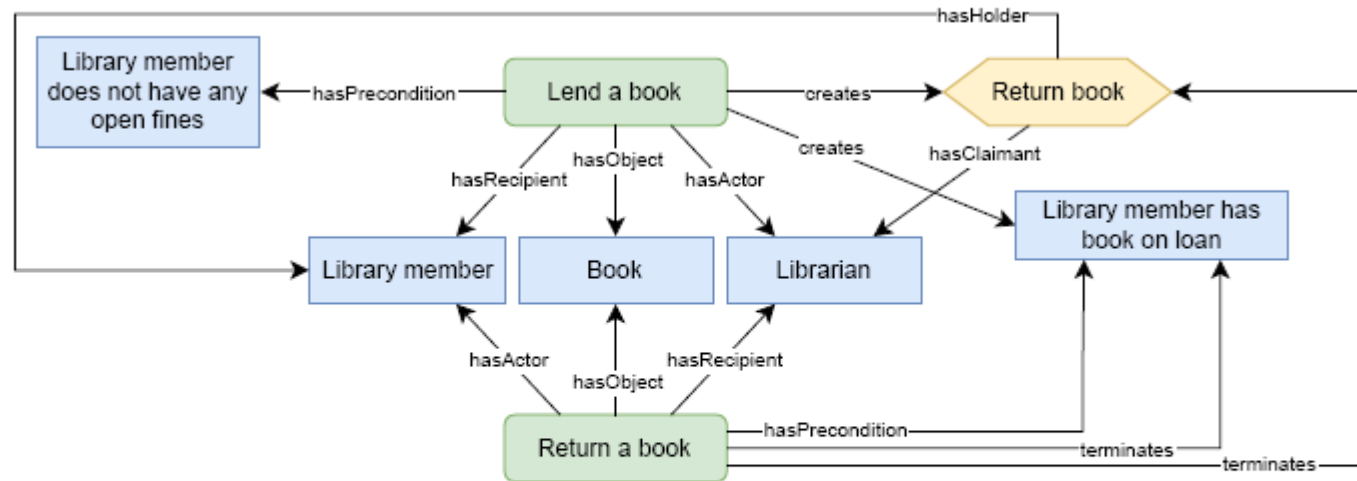
Figure 2. Representation in FLINT of the regulations for lending and returning books. Blue rectangles are instances of *AtomicFact*, green rounded rectangles are instances of *Act*, yellow hexagons are instances of *Duty*. Actions are omitted for readability.

Breteler et al. (2023)

Dataset

Requirements

- Questions must be simple (non-expert) ────────────────►

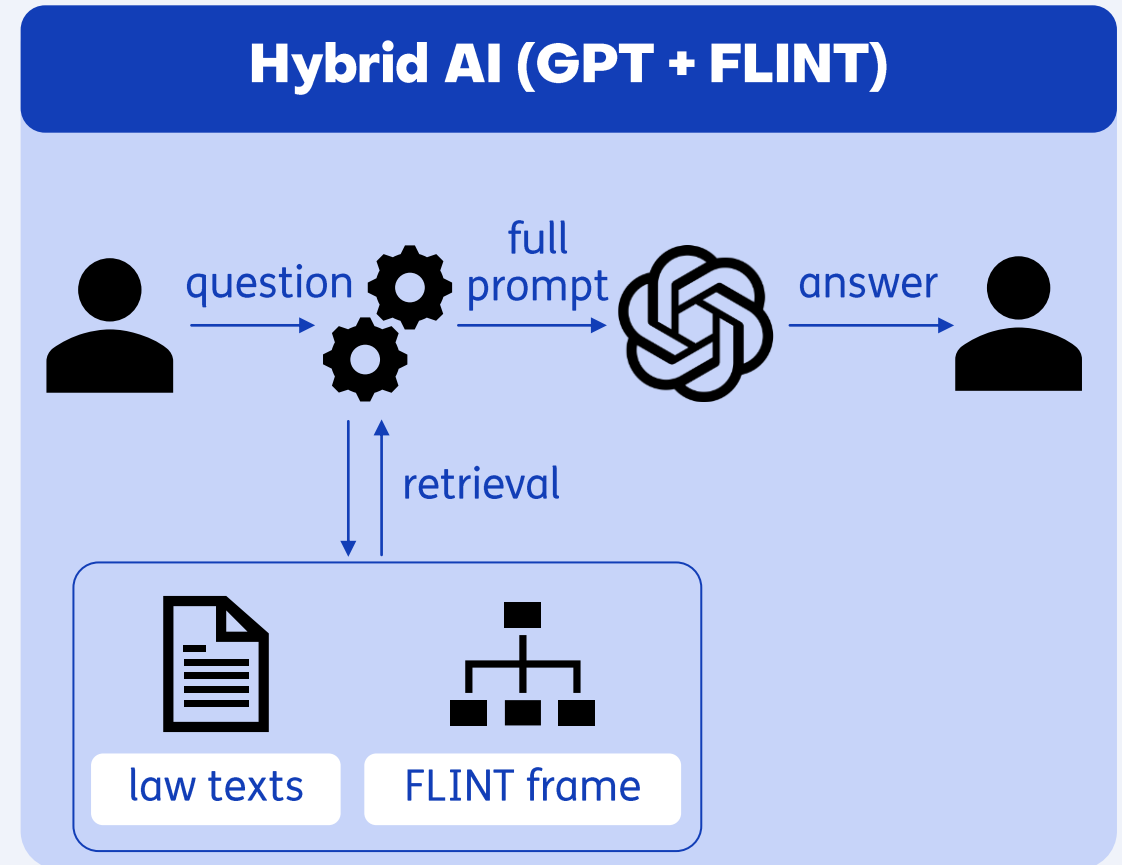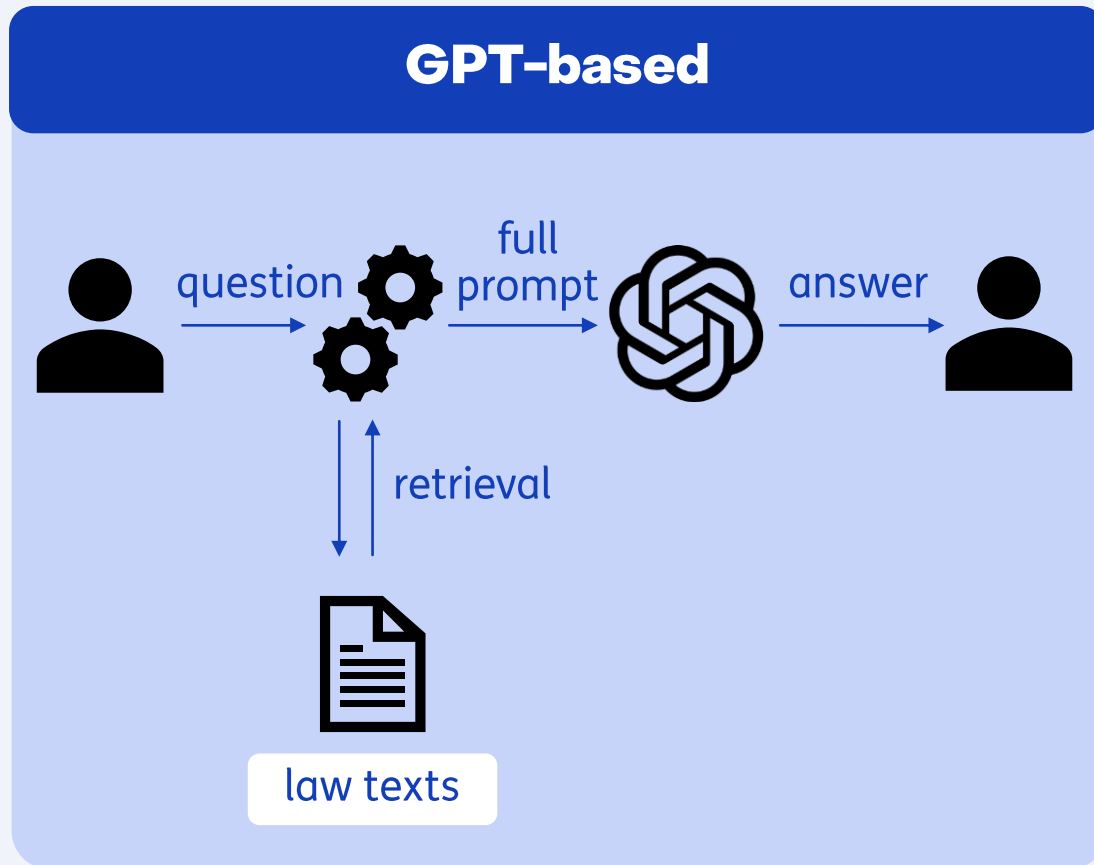- Answers must have a source (law and FLINT frame)

Can you help us find 'real' and relevant questions from non-experts?

Get in touch!

Questions - based on Aliens Act (Vreemdelingenwet) and Participation Act (Participatiewet)

- What is the maximum amount of equity I may have to qualify for a social assistance benefit (bijstand)?

- When should I apply for a return visa?

- My application for a residence permit has not been approved, what can I do now?

TNO innovation for life

Method

GPT-based

question → full prompt → answer

retrieval

law texts

Hybrid AI (GPT + FLINT)

question → full prompt → answer

retrieval

law texts   FLINT frame

TNO innovation for life

Thanks!

Get in touch with the Norm Engineering NLP team

Maaike de Boer
maaike.deboer@tno.nl

Roos Bakker
roos.bakker@tno.nl

Romy van Drie
romy.vandrie@tno.nl

Daan Di Scala
daan.discala@tno.nl

References

Bakker, R. M., M. H. T. de Boer, R. A. N. van Drie, and D. Vos (2022a). Extracting Structured Knowledge from Dutch Legal Texts: A Rule-based Approach. https://ceur-ws.org/Vol-3256/km4law1.pdf.

Bakker, R. M., R. A. N. van Drie, M. H. T. de Boer, R. van Doesburg, and T. M.van Engers (2022b). Semantic Role Labelling for Dutch Law Texts. https://aclanthology.org/2022.lrec-1.47.

Drie, R. A. N. van, M. H. T. de Boer, R. M. Bakker, I. Tolios, and D. Vos (2023). The Dutch Law as a Semantic Role Labeling Dataset. https://doi.org/10.1145/3594536.3595124.

Redelaar, F., R. Van Drie, S. Verberne, and M. De Boer (2024). Attributed Question Answering for Preconditions in the Dutch Law. https://aclanthology.org/2024.nllp-1.12.

Breteler, J., van Gessel, T., Biagioni, G., & van Doesburg, R. (2023). The FLINT Ontology: An Actor-Based Model of Legal Relations. In Knowledge Graphs: Semantics, Machine Learning, and Languages (pp. 227-234). IOS Press.

Harris, Z. (1954). Distributional structure. *Word*, 10(23): 146-162.

Mikolov, T., Chen, K., Corrado, G. S., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. arXiv preprint arXiv:1301.3781, 3781.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed Representations of Words and Phrases and their Compositionality. Advances in Neural Information Processing Systems, 26.

Devlin, J., Chang, M.-W., Lee, K., Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv:1810.04805v2

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems*, 33, 1877-1901.

TNO innovation for life