

September 27th, 2024

Bologna — HyperModeLex Seminar

LLMs for Unfairness Detection in Consumer Contracts

Marco Lippi

Department of Information Engineering

University of Florence

marco.lippi@unifi.it

Joint Work with...

LAW DEPT @EUI

CIRSFID @UNIBO



**Hans-Wolfgang
Micklitz**



**Francesca
Lagioia**



**Giovanni
Sartor**



**Giuseppe
Contissa**



**Ruta
Liepina**



**Przemyslaw
Palka**

LAW FACULTY
@JU KRAKOW



**Agnieszka
Jablonowska**

LAW FACULTY
@LEIDEN

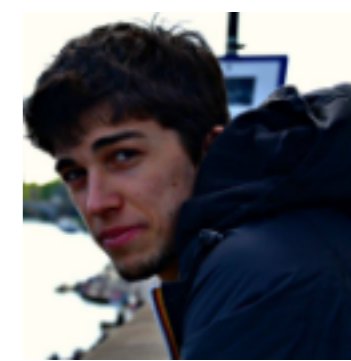


**Paolo
Torroni**



**Andrea
Galassi**

DISI @UNIBO



**Federico
Ruggeri**

CLAUDETTE

Automatically detect **potentially** unfair clauses in online contracts (a.k.a. Terms of Service): build an **automated CLAUse DETecTEr**

- Consumers agree but do not read
- NGOs have competence to control but lack resources
- Business keeps using unlawful clauses

<http://claudette.eui.eu>

CLAUDETTE

Annotation of ToS contracts according to eight categories for unfairness: unilateral termination, contract by using, arbitration, limitation of liability, ...

E.g., **unilateral termination** clauses give provider the right to suspend and/or terminate the service and/or the contract, and sometimes details the circumstances under which the provider claims to have a right to do so.

Academia.edu reserves the right, at its sole discretion, to discontinue or terminate the Site and Services and to terminate these Terms, at any time and without prior notice.

CLAUDETTE

A detection task, i.e., sentence classification, yet with many interesting research directions...

- Exploit background **knowledge** of the Law
- Provide **reasons** (explanations) behind unfairness
- Consider contracts in **multiple languages**

CLAUDETTE

- Lippi et al., “CLAUDETTE: an automated detector of potentially unfair clauses in online terms of service”, *Artificial Intelligence and Law*, 2019
- Galassi et al., “Cross-lingual Annotation Projection in Legal Texts”, *COLING 2020*
- Lippi et al., “The Force Awakens: Artificial Intelligence for Consumer Law”, *Journal of Artificial Intelligence Research*, 2020
- Jablonowska et al., “Assessing the Cross-Market Generalization Capability of the CLAUDETTE System”, *JURIX 2021*
- Ruggeri et al., “Detecting and explaining unfairness in consumer contracts through memory networks”, *Artificial Intelligence and Law*, 2022
- Galassi et al., “Unfair clause detection in terms of service across multiple languages”, *Artificial Intelligence and Law*, 2024

An Online Server

CLAUDETTE

An Automated Detector of Potentially Unfair Clauses

Copy your text here

Submit

[About](#) [Cite](#) [Contact](#)

<http://claudette.eui.eu/demo>

An Online Server

Potentially unfair clause #4

EXCEPT FOR CERTAIN TYPES OF DISPUTES MENTIONED IN THE ARBITRATION CLAUSE , YOU AND HEADSPACE AGREE THAT DISPUTES RELATING TO THESE TERMS OR YOUR USE OF THE PRODUCTS WILL BERESOLVED BY MANDATORY BINDING ARBITRATION , AND YOU WAIVE ANY RIGHT TO PARTICIPATE IN A CLASS-ACTION LAWSUIT OR CLASS-WIDE ARBITRATION .

Unfairness categories: **Arbitration**

[Hide/show rationales](#)

Potentially unfair clause #5

1.4 CHANGES TO TERMS Headspace reserves the right to change or update these Terms , or any other of our policies or practices , at any time , and will notify users by posting such changed or updated Terms on this page .

Unfairness categories: **Unilateral Change**

[Hide/show rationales](#)

The clause is potentially unfair for **Unilateral Change** since the provider has the right for unilateral change of the contract, services, goods, features for any reason at its full discretion, at any time (score = 0.834)

Potentially unfair clause #6

Your continued use of the Products constitutes your agreement to abide by the Terms as changed .

Unfairness categories: **Contract by Using**

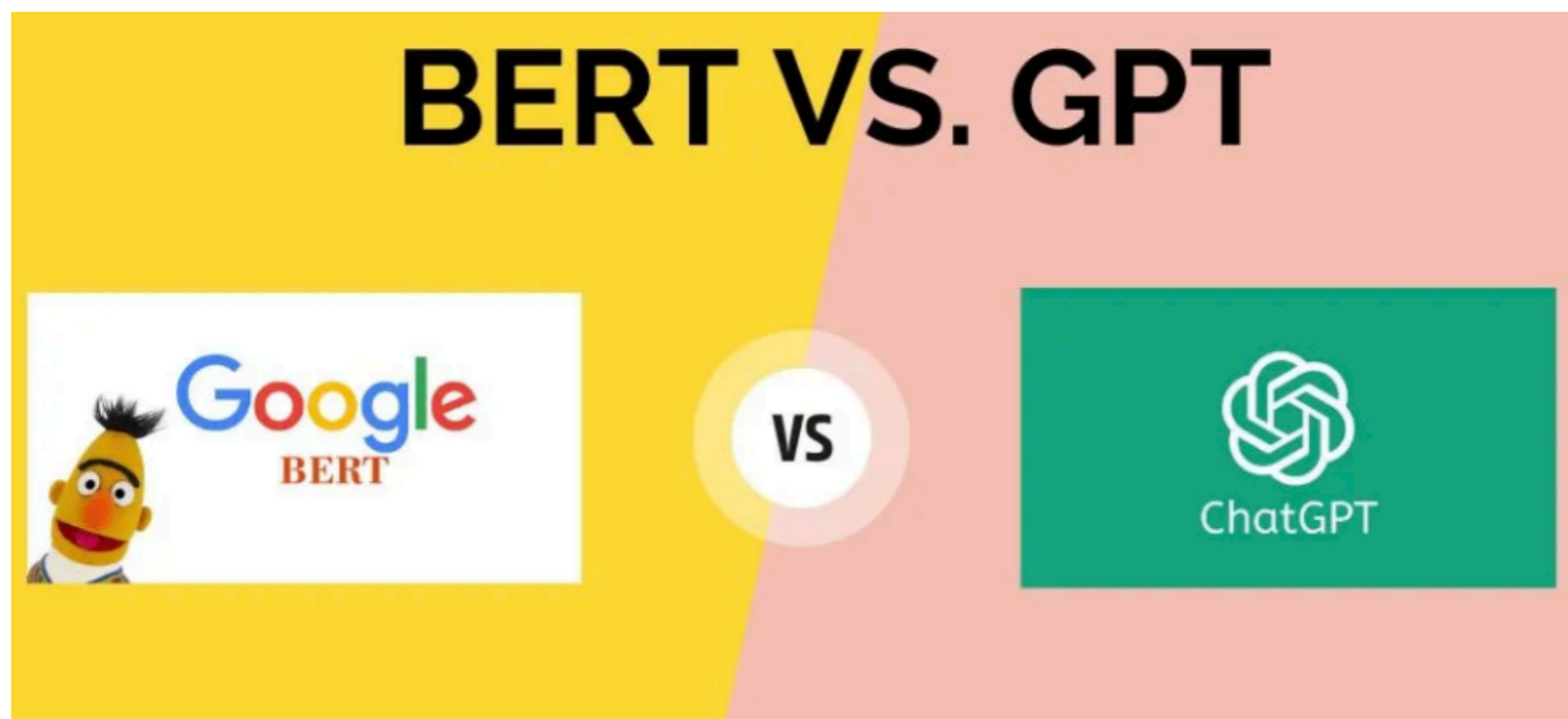
[Hide/show rationales](#)

<http://claudette.eui.eu/demo>

What about LLMs?

Can we do **better** using LLMs?

Or is this a **plain classification** task that is better to address with more traditional approaches?



How to exploit LLMs?

- Zero-shot vs. few-shot setting?
- Fine-tuning?
- Use rationales in the prompt?
- Use rationales in Retrieval Augmented Generation?
- Specialized (legal) LLM?
- Handling multilingualism?

Zero-shot

- One could investigate whether LLMs already have an internal knowledge of unfairness

```
marco@antares ~ % ollama run llama3.1:latest
```

```
>>> Please let me know whether this contract clause is unfair for the consumer: "Amazon reserves the right to refuse service, terminate accounts, terminate your rights to use Amazon Services, remove or edit content, or cancel orders in its sole discretion."
```

Zero-shot

A broad and seemingly draconian contract clause! While it's not uncommon for companies to reserve the right to take certain actions against customers who breach their terms of service, this particular clause is indeed concerning from a consumer protection standpoint. Here's why:

1. ****Lack of specificity****: The clause doesn't provide clear guidelines on what constitutes "sufficient" behavior or circumstances that would trigger such actions. This makes it difficult for consumers to understand their rights and responsibilities.
2. ****Overly broad discretion****: Granting Amazon the right to act in its "sole discretion" gives the company too much latitude, potentially allowing them to arbitrarily take action against customers without any clear justification.
3. ****No recourse or appeals process****: The clause doesn't specify how a consumer can appeal Amazon's decisions if they feel they've been unfairly penalized.
4. ****Potentially excessive and unfair consequences****: Terminating an account, removing content, canceling orders, or refusing service without notice or explanation can have significant financial and reputational implications for consumers.

Zero-shot

From a consumer protection perspective, this clause could be seen as unfair because:

1. ****It creates uncertainty and anxiety****: Consumers may feel uncertain about their rights and responsibilities when interacting with Amazon, making it difficult to trust the platform.
2. ****It restricts access to essential goods or services****: In some cases, refusing service, terminating accounts, or canceling orders might limit consumers' ability to obtain necessary products or services.

To put this into perspective:

- * The Consumer Protection from Unfair Trading Regulations 2008 (UK) and similar regulations in other countries require companies to be transparent about their policies and procedures.
- * The European Union's General Data Protection Regulation (GDPR) emphasizes the need for fairness, transparency, and accountability in data processing activities.

Zero-shot

In light of these considerations, a more consumer-friendly approach might involve revising this clause to:

1. ****Clearly outline grounds for termination or account suspension****
2. ****Provide consumers with reasonable notice and opportunities to appeal Amazon's decisions****
3. ****Specify circumstances under which Amazon will take specific actions (e.g., order cancellation, content removal)****

Please note that contract terms can vary significantly depending on jurisdiction and local laws, so it's always a good idea to consult relevant regulations and seek professional advice if you're unsure about the fairness of a particular clause.

If you'd like me to suggest possible revisions or provide further context, feel free to ask!

Few-shot

- We may add examples in the prompt

```
marco@antares ~ % ollama run llama3.1:latest
```

```
>>> This is an example of a contract clause that is potentially unfair for the consumer according to the European Law: "We can also terminate or change the Service, remove or block content or information shared on our Service, or stop providing all or part of the Service if we determine that doing so is reasonably necessary to avoid or mitigate adverse legal or regulatory impacts on us.". Can you tell me whether the following clause now is potentially unfair for the consumer: "We may modify, suspend, or terminate your access to or use of our Services anytime for any reason, such as if you violate the letter or spirit of our Terms or create harm, risk, or possible legal exposure for us, our users, or others."
```

Few-shot

While this clause shares similarities with the first one in terms of potentially allowing a service provider to terminate or modify services without clear justification, it differs significantly in its specificity and conditions.

Here are some reasons why:

1. ****Specificity****: The second clause clearly states that termination or modification can happen if the consumer "violate[s] the letter or spirit of our Terms" or if they create any form of harm or risk for others. This gives a clear understanding to consumers about under what circumstances their access to services might be terminated.
2. ****Conditions****: Unlike the first clause, which seems quite broad in its application ("adverse legal or regulatory impacts"), this one specifies certain actions that would lead to termination or modification (violating terms, creating harm).
3. ****Transparency****: The language used here is more direct and less ambiguous compared to the first clause.

Few-shot

However, there are still some potential issues with fairness for consumers:

- * **Subjectivity**: The term "spirit of our Terms" might be open to interpretation, which could lead to disputes over whether a consumer has indeed violated this part.
- * **Lack of clear notice**: While the clause specifies reasons for termination or modification, it doesn't necessarily outline what steps will be taken before such actions are implemented. This could leave consumers feeling they've been unfairly targeted without adequate warning.

Overall, while less ambiguous than the first example, there are still points to consider regarding fairness and consumer protection in this clause.

Few-shot vs. Fine-tuning

- Few-shot learning still works in-context
- One could fine-tune the LLM (with gradient descent) using a training set of labeled examples
- Much more time-consuming!!!
- Maybe use an LLM already specialized on legal text

Using rationales

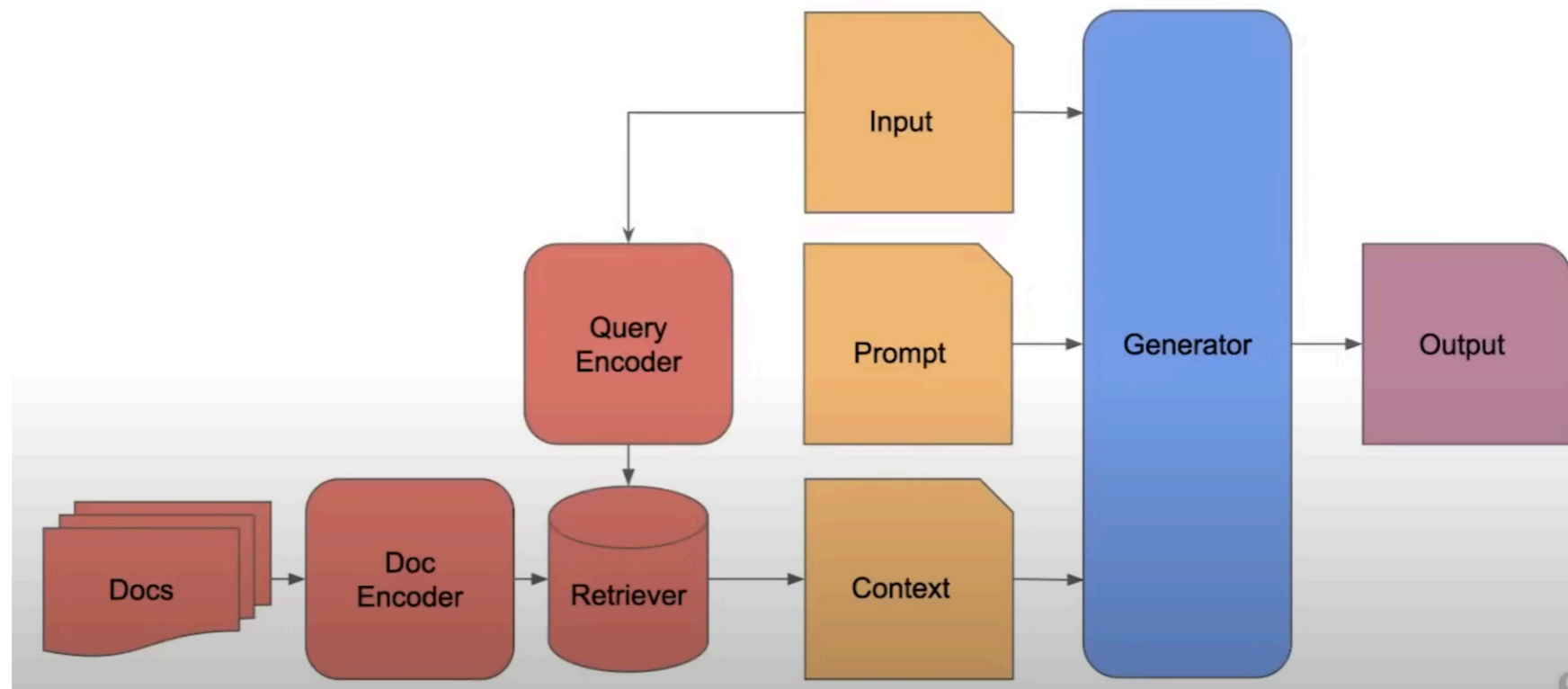
- Legal rationales are **explanations**, provided by legal experts, that motivate the unfairness of a clause

“With new products, services, and features launching all the time, we need the flexibility to make changes, impose limits, and occasionally suspend or terminate certain offerings.” — Endomondo ToS, 2016

LEGAL RATIONALE: the clause is potentially unfair since the provider has the right for unilateral change of the contract/services/goods/features to maintain a level of flexibility to amend and update services, including discontinuation

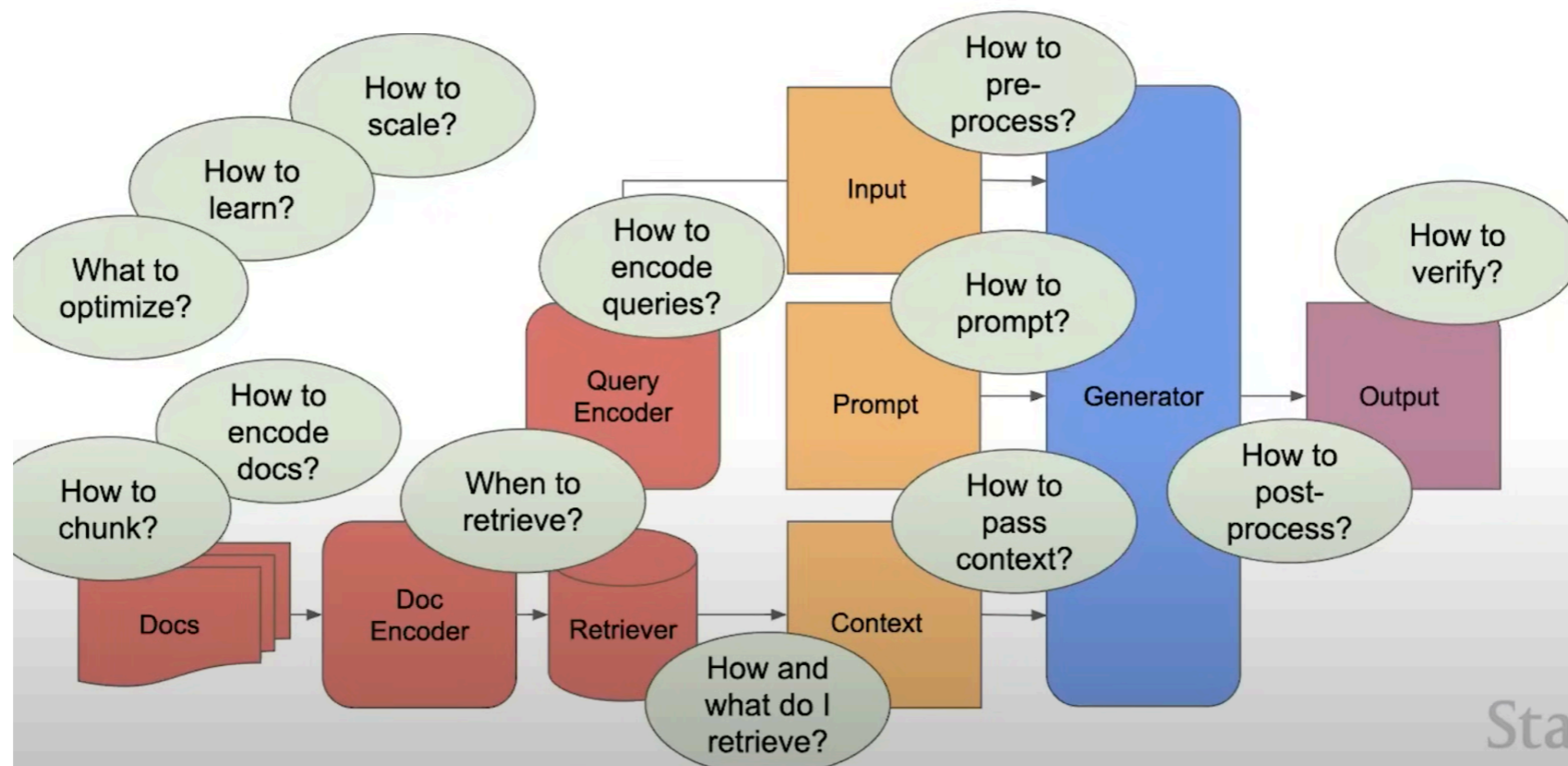
Using rationales

- Use them in the prompt
- Use them in Retrieval Augmented Generation



Using rationales

- Use them in the prompt
- Use them in Retrieval Augmented Generation

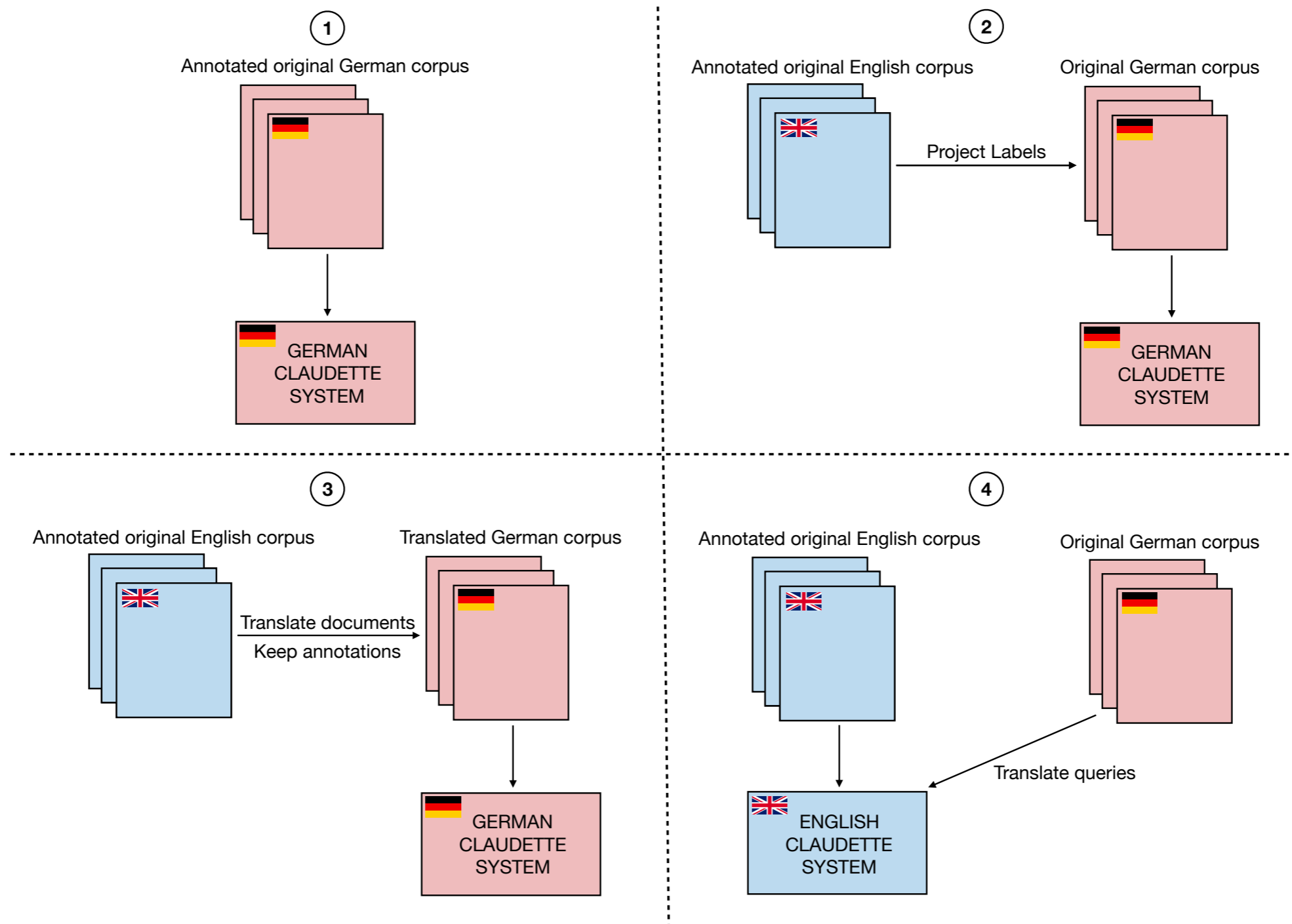


Multilingualism

We compared several methodologies in order to make **CLAUDETTE** a **multilingual system**

- Building novel corpus for each language
- Annotation projection across languages
- Automatic translation of training set
- Automatic translation of test set

Multilingualism



Future works

Make an experimental **comparison** of approaches

- Zero-shot vs. Few-shot learning
- Prompt engineering
- Fine-tuning or specialized LLM
- Using legal rationales
- Multilingual data