

DIREGA

Denkfabrik: "Legal Tech" am 25.10.2024

- ❖ Einführung (Adrian)
- ❖ Datenlage & Data Mining (Maier, Prettnner)
- ❖ Computerlinguistische Analyse und Auswertung von Zwischenverfügungen (Evert)
- ❖ Datensuche und -extraktion (Evert)
- ❖ Aufbereitung des juristischen Hintergrundwissens (Stürmer)
- ❖ Formalisierungsprozess (Kohlhase, Schröder)

Einführung

Ziel, Projekt, Team

Viele setzen große Hoffnungen in die *Leistungsfähigkeit* von KI-Verfahren zur Unterstützung bei juristischer Arbeit.

Wir wollen wissen ...

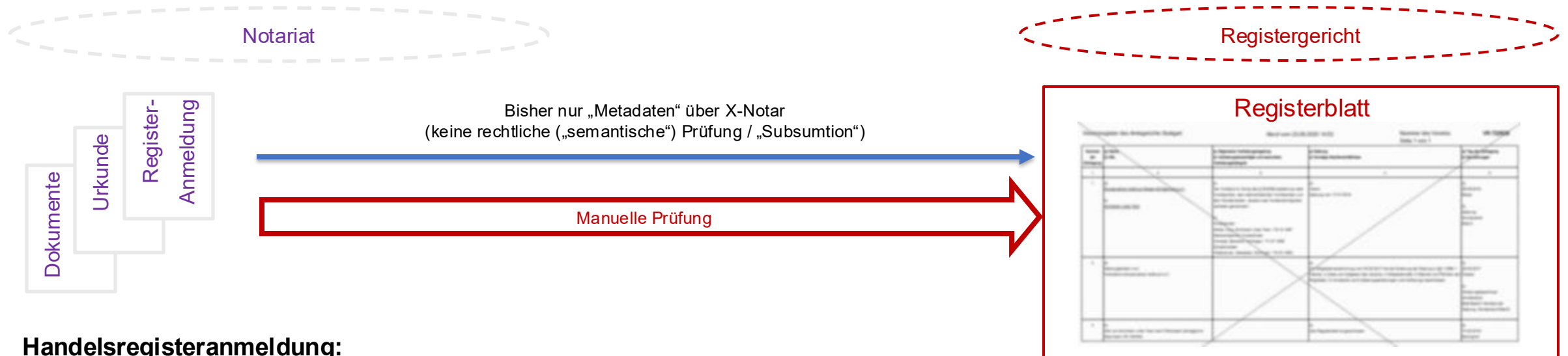
wie *leistungsfähig* und hilfreich können KI-Verfahren für Juristinnen und Juristen wirklich sein?

Projekt im zeitlichen Verlauf



Milestones

| | |
|----------------------------|--------------------------------------------------------------------------------------------------------------------|
| 19.10.2020 | Erste Vorstellung der Projektidee in München beim IT-Rat der LNotK |
| 01.10.2021 | Darlegung Projektidee in der Generalversammlung der BNotK in Travemünde zur Beschlussfassung |
| 25.08.2023 | Unterzeichnung des Forschungsvertrages (BNotK, FAU, BaySTMJ, LNotK, Adrian) |
| 01.10.2023 | Stellenausschreibungen und Einstellung der wissenschaftlichen Mitarbeiter |
| 01.11.2023 | Kontaktaufnahme mit BayStMJ zur Beschaffung der veröffentlichten und nicht-veröffentlichten Daten |
| 12/2023 bis 02/2024 | Organisatorisches, Räume an der FAU, Erstellung einer Website, etc. |
| 03/2024 | Offizieller Projektbeginn |
| 14.03.2024 | Einrichtung großer Jourfix mit allen Wissenschaftlern je Donnerstags 10-12 Uhr |
| 19.03.2024 | Feierliches Kick-Off Meeting des Projekts DIREGA an der FAU im Wassersaal der Orangerie im Schlosspark in Erlangen |
| 25.10.2024 | Denkfabrik Legal-Tech |
| 03/2027 | Abschluss des Projektes und Vorstellung der Ergebnisse |
| Ab 2029 | Der geburtenstärkste Jahrgang von 1964 geht in Rente |



Handelsregisteranmeldung:

angemeldet, was folgt:

-

Zum neuen Geschäftsführer wurde bestellt:
Herr Maximus Mustermann, geb. am 01.01.20XX, wohnhaft in
Musterstadt.

Er ist als Geschäftsführer -auch bei Vorhandensein von weiteren
Geschäfts-führern- stets einzeln zur Vertretung der Gesellschaft
berechtigt und von den Beschränkungen des § 181 BGB befreit.

-

Ich, Herr Maximus Mustermann, versichere, dass

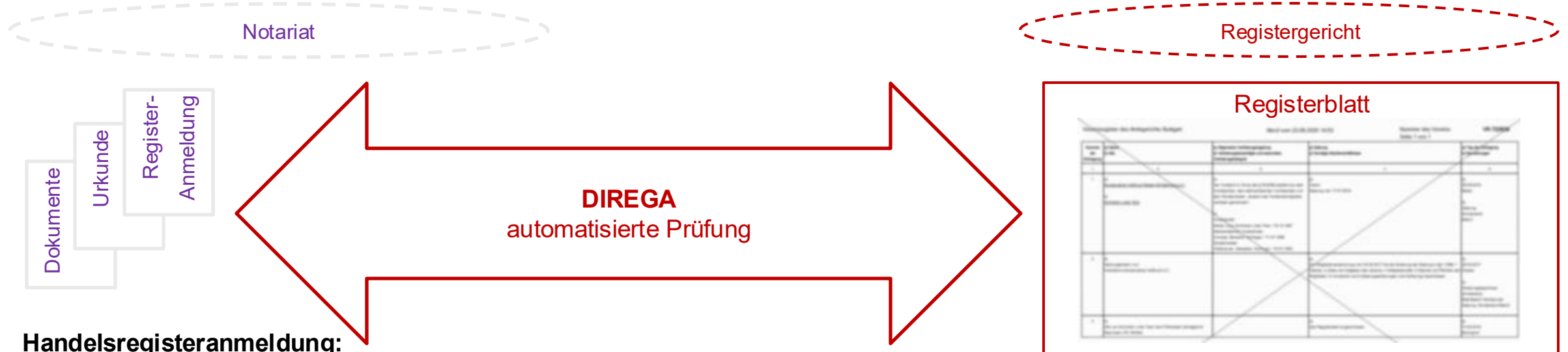
Handelsregistereintrag:

„b)

Geschäftsführer:

Mustermann, Maximus, Musterstadt,
*01.01.20XX

einzelvertretungsberechtigt; mit der Befugnis,
im Namen der Gesellschaft mit sich im
eigenen Namen oder als Vertreter eines
Dritten Rechtsgeschäfte abzuschließen.“



Handelsregisteranmeldung:

angemeldet, was folgt:

-

Zum neuen Geschäftsführer wurde bestellt:

Herr Maximus Mustermann, geb. am 01.01.20XX, wohnhaft in Musterstadt.

Er ist als Geschäftsführer -auch bei Vorhandensein von weiteren Geschäftsführern- stets einzeln zur Vertretung der Gesellschaft berechtigt und von den Beschränkungen des § 181 BGB befreit.

-

Ich, Herr Maximus Mustermann, versichere, dass

Handelsregistereintrag:

„b)

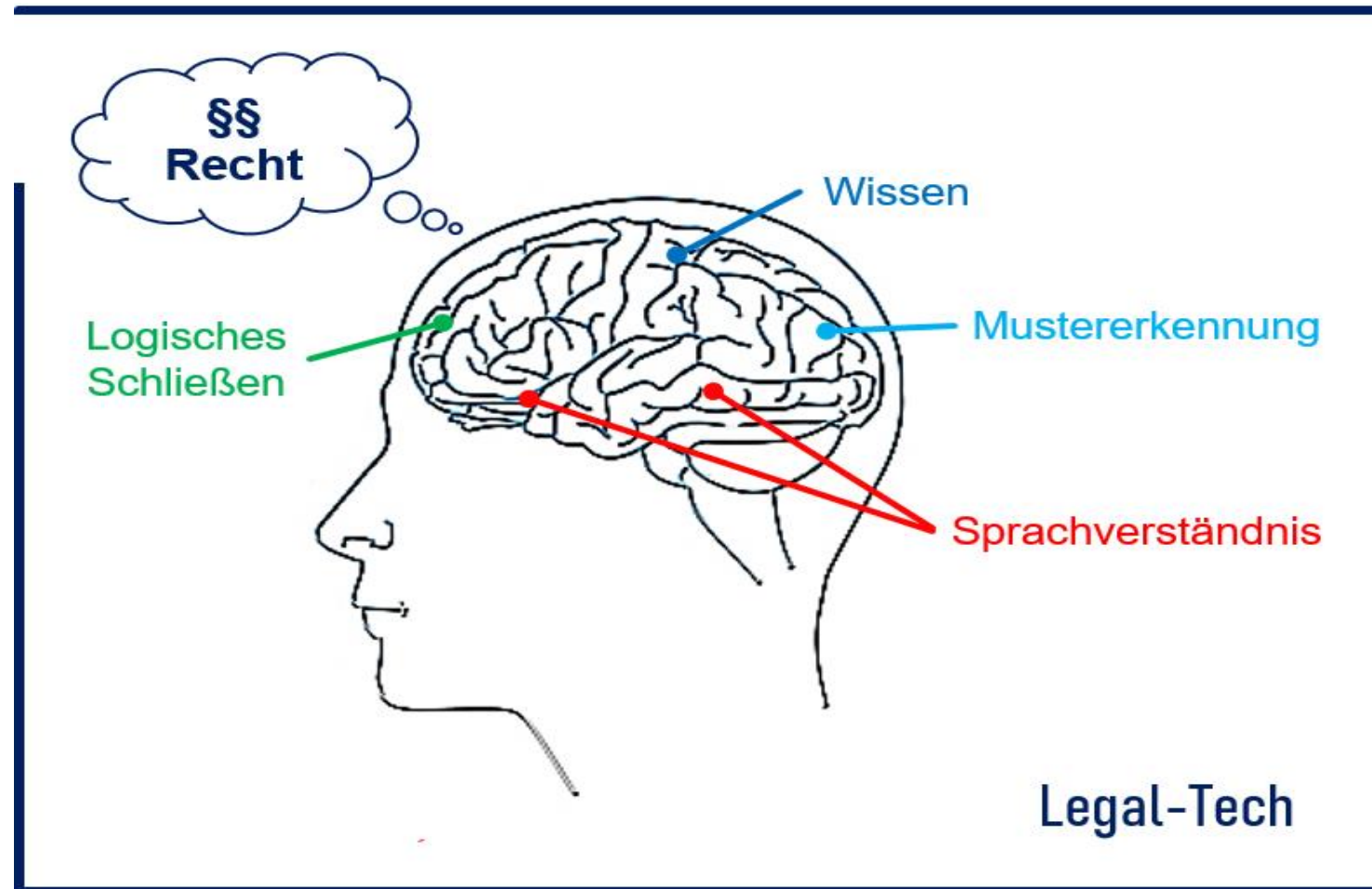
Geschäftsführer:

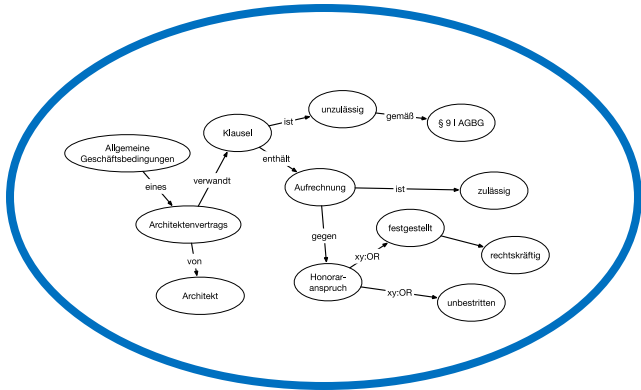
Mustermann, Maximus, Musterstadt,
*01.01.20XX

einzelvertretungsberechtigt; mit der Befugnis,
im Namen der Gesellschaft mit sich im
eigenen Namen oder als Vertreter eines
Dritten Rechtsgeschäfte abzuschließen.“

Projektidee

Analogie: Gehirn und KI-Systeme





Symbolische
und
subsymbolische
KI

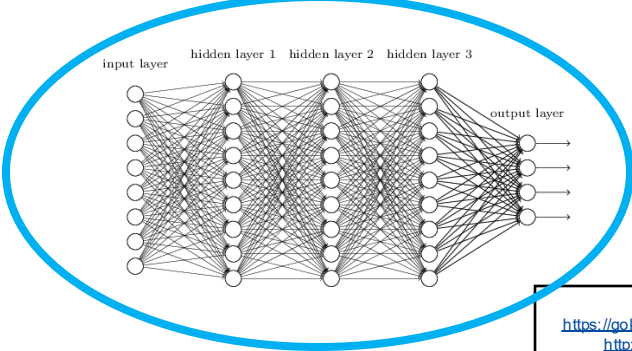
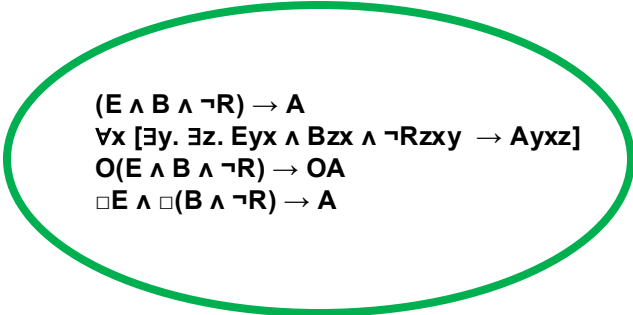
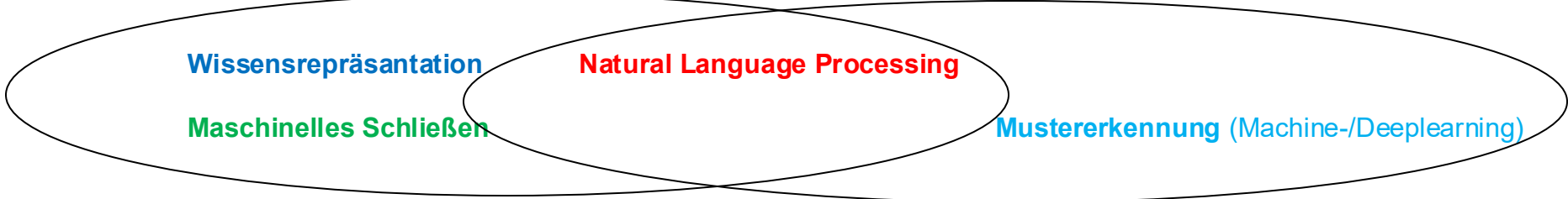
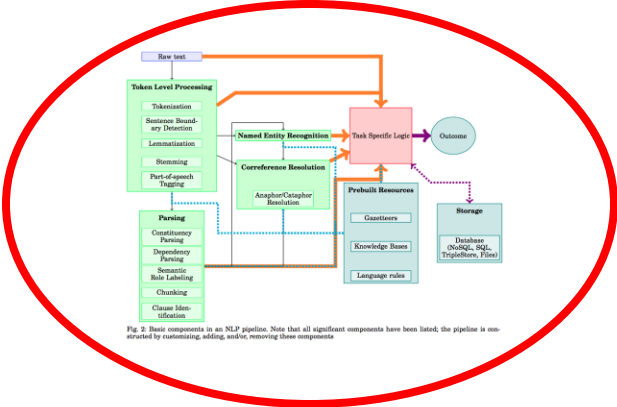


Abbildung:
<https://gokulchittaranjan.files.wordpress.com/2015/09/pipeline1.png>
<http://neuralnetworksanddeeplearning.com/images/tikz36.png>
Alexey Cheptsov (HLRZ Stuttgart)
Eigene Darstellung

Projektteam



Institutionen, Principal Investigators, Wissenschaftliche Mitarbeiter



Prof. Dr. Axel Adrian
(Wissenschafts- und Rechtstheorie)

Osman Anil Basaran, Ass. Jur LL.M. Legal Tech (*Regensburg*)
Michael Gritz, Notarassessor
Dr. Verena Stürmer, Notarassessorin

Legal



Prof. Dr. Stephanie Evert
(Natural Language Processing)

Steffen Bothe, M.A.
Nathan Dykes, M.A.

CCL



Prof. Dr. Michael Kohlhase
(Wissensrepräsentation)

Max Rapp, M.Sc. (phil.), Ms.Sc. (logic)

KWARC



Prof. Dr. Andreas Maier
(Mustererkennung)

Stephan Prettner, M.Sc.
Sven Grünke, IT & Lab
Sebastian Wind, M.Sc.

ML



Prof. Dr. Lutz Schröder
(Theoretische Informatik)

Merlin Humml, M.Sc.
Johannes Lindner (stud. B.Sc.), studentische Hilfskraft
Moritz Blöcher, B.Sc. (stud. M.Sc.), Masterand

T.SC

Sonstige Beteiligte

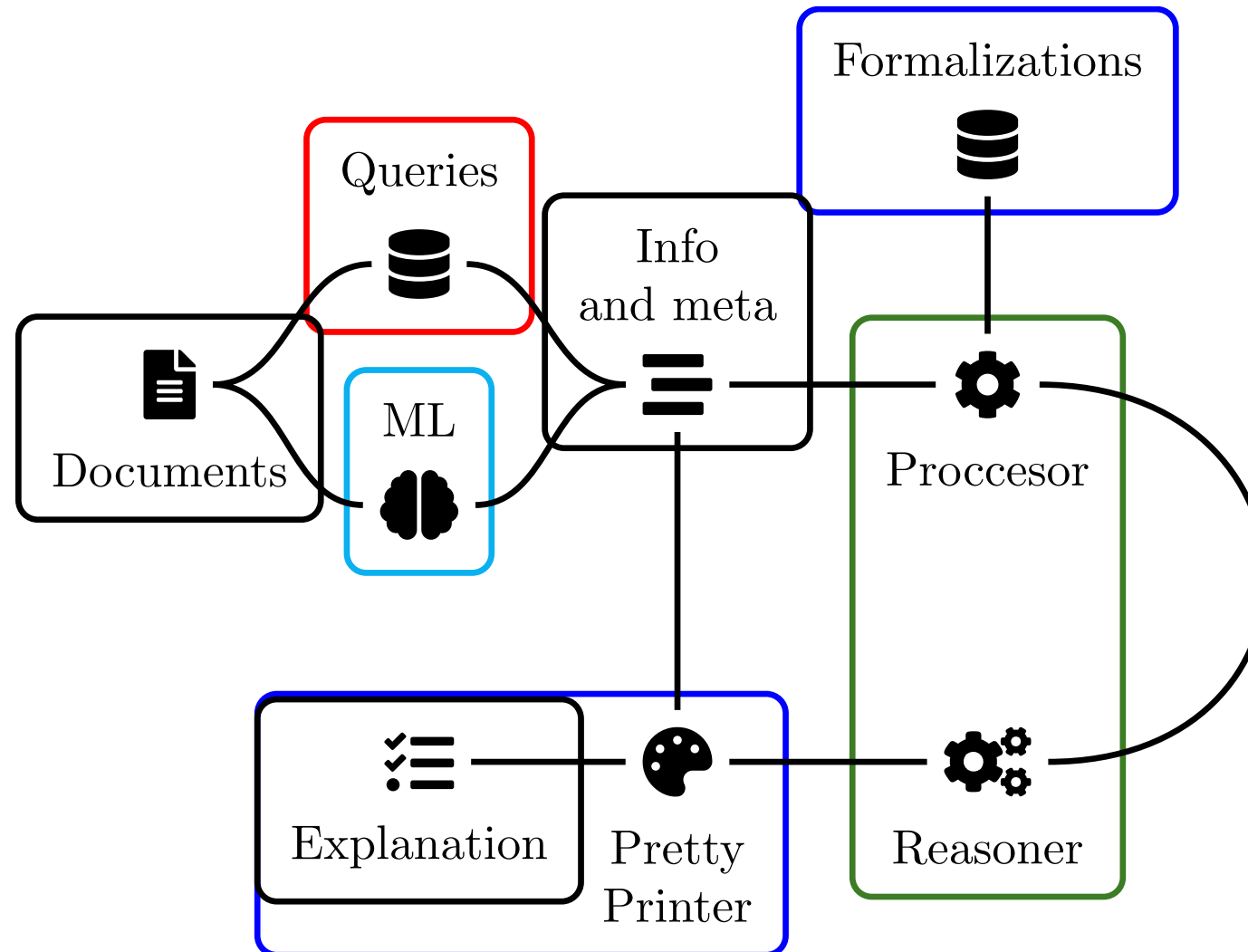
Prof. Dr. Christian Bergler (Mustererkennung)
Lucca Baumgärtner, B.Sc., stud (M.Sc.)
Aurelius Adrian, stud. (B.Sc.)

Im Projekt DIREGA soll ...

- ❖ mit den - hinsichtlich Umfang und Vollständigkeit einzigartigen - Daten die seit 2007 zwischen Notariaten und Registergerichten im Wege des elektronischen Rechtsverkehrs (XNP) bei Handelsregisteranmeldungen ausgetauscht werden
 - ❖ Grundlagenforschung betrieben werden und
 - ❖ ein Prototyp entwickelt und evaluiert werden, der
 - ❖ mit symbolischen, sub-symbolischen und auch hybriden KI-Systemen
 - ❖ Notariate (und ggf. auch spiegelbildlich Registergerichte)
 - ❖ bei Entwurf, Vollzug und Kontrolle von Handelsregistervorgängen
 - ❖ möglichst automatisch unterstützt.
-
- ❖ So können Erkenntnisse auch ganz allgemein für KI-Anwendungen zur Modellierung juristischer Entscheidungen am Beispiel der Registervorgänge gewonnen werden.
-
- ❖ Besonders im Fokus steht dabei von Anfang an die Erklärbarkeit und die Verständlichkeit sowohl des maschinell repräsentierten Wissens als auch der generierten Entscheidungsvorschläge

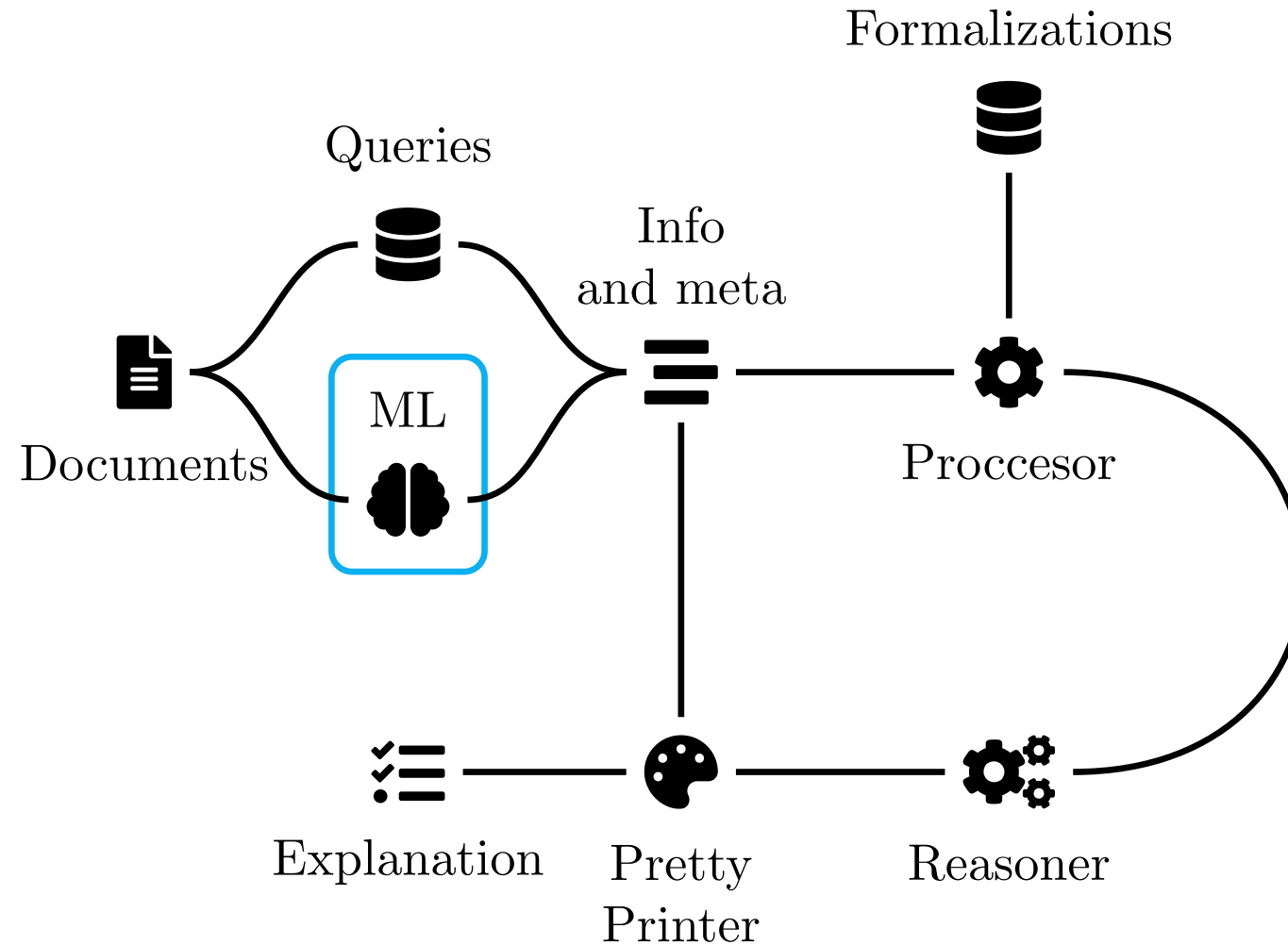
Projekt und Ziele

Pipeline des Registerassistenten



Datenlage & Data Mining

Vorverarbeitung, Dokumentenanalyse (insb. Zwischenverfügungen), Data Mining i.e.S.



Vorbereitung:

- ❖ Abholung zweier mit Rohdaten befüllten Festplatten in München
- ❖ Analyse und Verarbeitung der enthaltenen Daten
- ❖ Verwendung für diverse Anwendungszwecke innerhalb des Projekt-Scopes

Einrichtung der Hardwarelösung:

- ❖ Verschlüsselung des Systems und der verwendeten Datenspeicher
- ❖ Integration benötigter Systemsoftware
- ❖ Konsultation und Beschaffung der benötigten Datenbanklizenz
- ❖ Export und Transport der Testdaten des Handelsregisters
- ❖ Anpassung der Systemparameter des Zielsystems an die Gegebenheiten des Handelsregisters
- ❖ Import der Informationen in die Datenbank (Massendaten)
- ❖ Integration des Dokumentenmanagementsystems

Ergebnis: Lokale Datenbank

- ❖ Enthält frühere Handelsregisteranmeldungen mit den dazugehörigen Dokumenten und Entscheidungen der deutschen Registergerichte in Bayern
- ❖ Ausgangspunkt für Analyse weiteren Projektfortschritt

Datenmaterial (Σ):

- ❖ Komprimierter SQL-Dump (Oracle-Datenbank) umfasst 1,3 TB (unkomprimiert: 5 TB)
- ❖ 11.917.322 Dateien verschiedener Dateitypen (Bilder, Dokumente, ...)
- ❖ Strukturierte Dateninformationen:
 - 1.219.814 Fälle für verschiedene Handelsregistereinträge
 - 445.023 Handelsregistereinträge (Adressen, Rechtsformen, Firmensitz, ...)
 - 30 Rechtsformen (Einzelfirmen, Personengesellschaften, Kapitalgesellschaften, ...)
 - 96 Dokumentenarten (Interne Mitteilungen, Anträge an Registergerichte, Vollstreckungsdokumente, ...)
 - 124 Fallarten (Neueintragungen, Löschungen, Gesellschafterlisten, Jahresabschlüsse)

- ❖ **Klassifikation** von übermittelten **Dokumenten**
 - Handelt es sich um das korrekte Dokument?
 - Anstoßen interner Prozesse anhand des eingereichten Dokumentes
- ❖ **Extraktion** von spezifizierten **Informationen**
 - Wer ist der Geschäftsführer?
 - Bis wann muss ein bestimmtes Problem behoben sein?
- ❖ Erstellung und Durchsuchung automatisch erstellter **Dokumentzusammenfassungen**
- ❖ Identifizierung von spezifischen **Entitäten** (Personen, Organisationen, Orten, ...)
- ❖ Optische Zeichenerkennung (OCR)
 - Analyse der eingereichten Dokumente
 - Analyse verwendeter Unterschriften
- ❖ Sentimentanalyse
- ❖ **Konsistenzprüfung** erstellter und eingereicherter Dokumente
 - Existieren die benötigten Informationen?

Erstellter Datensatz:

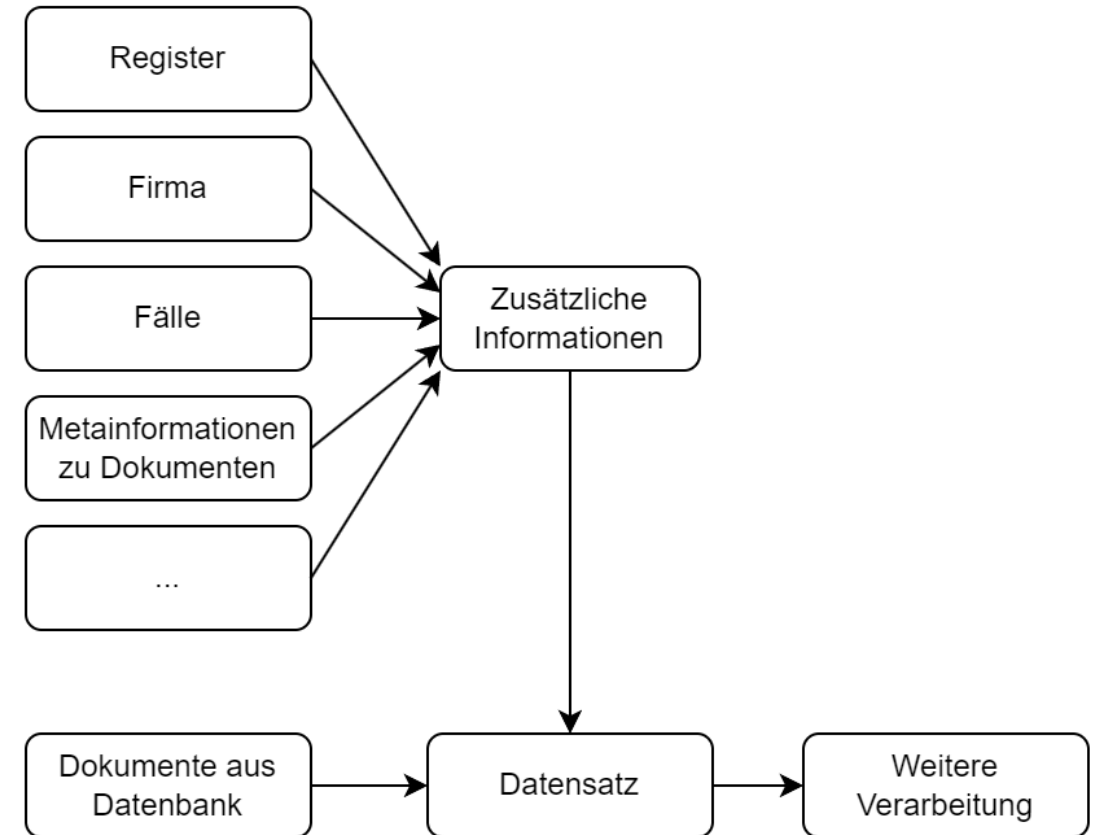
- ❖ Dokumente (Gesellschafterlisten, Satzungen, ...)
- ❖ Zusätzliche Informationen zur:
 - Klassifikation, Informationsextraktion, ...

Datenvorbereitung (auszugsweise):

- ❖ Zwischenverfügungen vereinheitlicht
- ❖ Dateien entfernt (Ungültige Dateitypen, nicht entpackbare Archive, ...)
- ❖ Extraktion der Inhalte aus den Zwischenverfügungen

Ergebnis:

- ❖ Informationen zu den einzelnen Dokumenten zur weiteren Verarbeitung



Ziel:

❖ Klassifikation eines Dokumentes

- Eingabe: Dokument
- Ausgabe: Satzung, Gesellschafterliste, ...

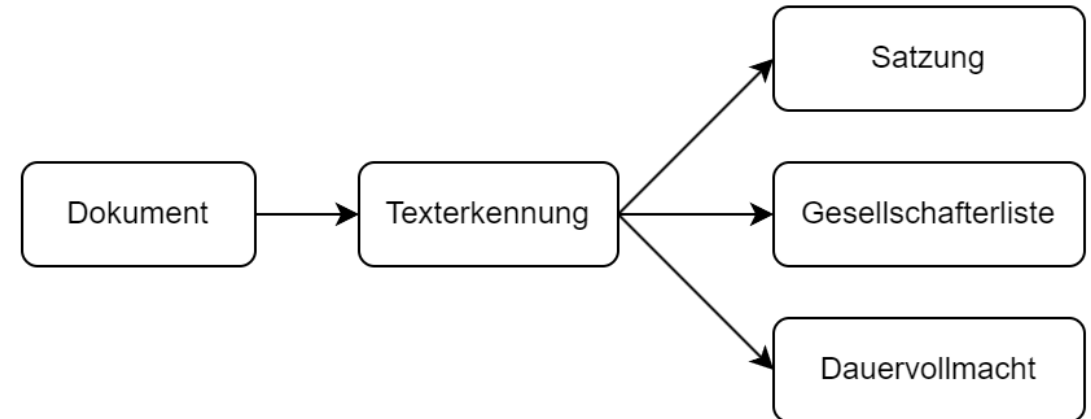
Voraussetzungen:

❖ Dokumente

- ✓ Unterschiedliche Dokumententypen

❖ Korrekte Klassifikationen (Supervised Learning)

- ✓ Metadaten in der Datenbank

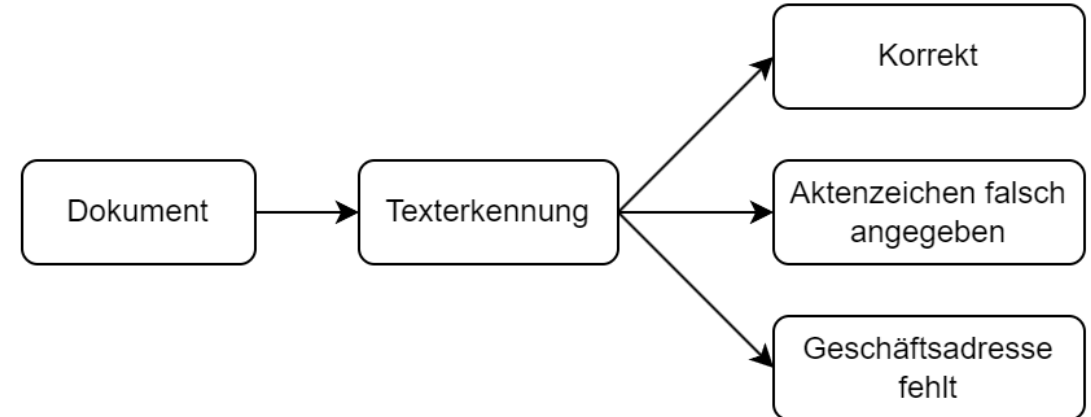


Ziel:

- ❖ Analyse eingereichter Dokumente
 - Eingabe: Dokument
 - Ausgabe: Nicht gültig, weil <Grund>

Voraussetzungen:

- ❖ Dokumente
 - ✓ Viele Dokumente
- ❖ Gründe, warum Dokumente ungültig sind
 - ✗ Gründe sind nicht **direkt** hinterlegt
 - **Wie kann man dieses Problem trotzdem lösen?**



Ist - Zustand:

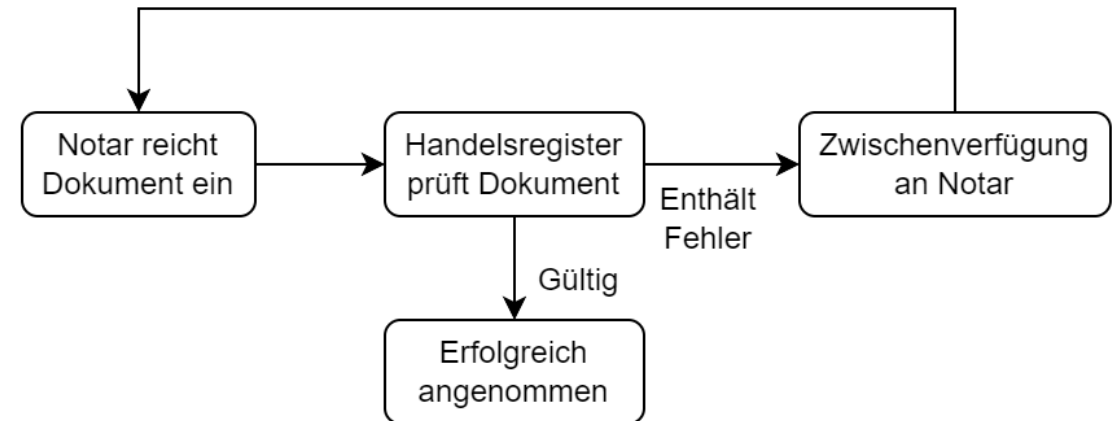
- ❖ Viele zu analysierende Dokumente vorhanden
- ❖ Keine Felder wie: <Dokument X hat Problem Y>

Soll - Zustand:

- ❖ Warum wurde Dokument X abgelehnt?
- ❖ Was muss für eine Annahme durchgeführt werden?

Idee:

- ❖ Zwischenverfügungen enthalten Informationen über Probleme
- ❖ Extraktion dieser Informationen und Zuordnung zu den Dokumenten



Zwischenverfügungen:

- ❖ enthalten Informationen über Probleme

Aufbau:

- ❖ Header, Footer
- ❖ Informationen:
 - Datum
 - Bearbeiter
 - Handelsregisternummer
 - **Tatsächlicher Inhalt**
 - ...
- ❖ **Optimalerweise:**
 - Warum wurde Dokument X abgelehnt?
 - Was muss für eine Annahme durchgeführt werden?

Fax: 012/3456-789

Amtsgericht München, 80325 München

Herrn Notar
Max Mustermann
Marienplatz 79
80213 München

Bei Rückfragen wenden Sie sich bitte an:
Telefon: 012/3456-789

Allgemeine Öffnungszeiten:
Mo-Do 8.30-11.30 und 13.00-15.00 Uhr
(Einsicht durchgehend); Fr 8.30-12.00 Uhr
wegen gleitender Arbeitszeit individuelle
Terminvereinbarung möglich

Öffentliche Verkehrsmittel:
Straßenbahnlinien 20/21, Haltestelle Lothstraße
Straßenbahnlinie 12, Haltestelle Infanteriestraße
Buslinie 53, Haltestelle Infanteriestraße
Buslinie 154, Haltestelle Infanteriestraße Süd

Online-Einsicht:
www.handelsregister.de

Ihre Nachricht vom - Ihr Zeichen

Bei Antwort bitte angeben:
Unsere Geschäftsnummer
HRB 318508 (Fall 29)

Datum
12.10.2010

Herbert Fuchs AG, Sitz: Starnberg

Sehr geehrter Herr Notar Mustermann,

die vorgelegte Anmeldung kann derzeit leider nicht vollzogen werden. Es steht folgendes Vollzugshindernis entgegen:

Im Rahmen der Neufassung der Satzung wurden nicht nur der Sitz und der Unternehmensgegenstand, sondern auch die allgemeine Vertretungsregelung geändert. Dies ist ausdrücklich anzumelden.

Es muss mit Zurückweisung der Anmeldung gerechnet werden, wenn den gerichtlichen Anforderungen nicht bis zum 31.01.2011 entsprochen wird.

Mit freundlichen Grüßen

Was ist ein Sprachmodell (Large Language Model, LLM)?

- ❖ Darauf ausgelegt, (menschliche) Sprache zu verstehen
- ❖ Texte analysieren, verstehen, verarbeiten
- ❖ Einsatz zum Beispiel:
 - Übersetzungen, Chatbots, Sentiment-Analyse, ...
- ❖ Chat**GPT** (OpenAI), BERT (Google), LLaMA (Meta)
- ❖ Basieren auf der sog. Transformerarchitektur

Möglichkeit zur Interaktion:

- ❖ LLMs können auf Anfragen (sog. "Prompts") antworten
- ❖ Dies kann man zur automatisierten Bearbeitung von Texten nutzen (mithilfe von "Prompt Engineering")
 - Extraktion existierender Probleme aus Zwischenverfügungen

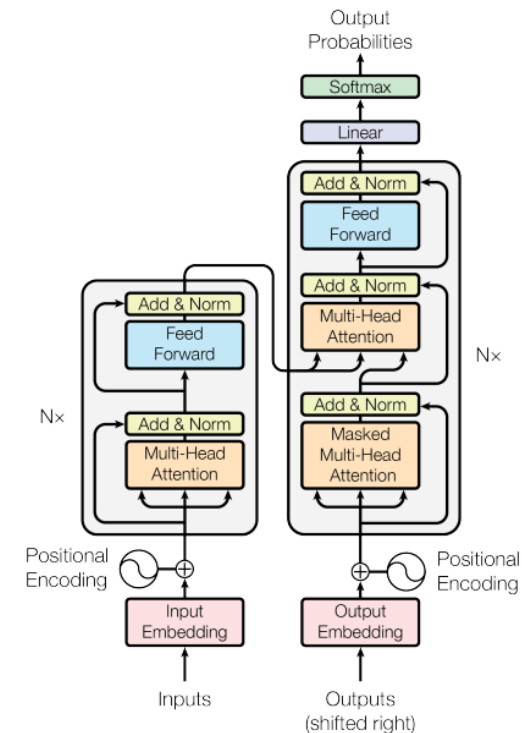


Abbildung: Visualisierung der Transformerarchitektur.
Quelle: Vaswani et al. (2017, S. 3)

Was ist Prompt Engineering?

- ❖ Optimierung eines Prompts, um eine bessere Ausgabe zu erhalten

- ❖ Unser Beispiel:

- Eingabe: Text der Zwischenverfügung
- Ausgabe: Relevante Informationen
 - Probleme, Lösungsvorschläge, ...

- ❖ Verwendete Techniken:

- Zero-Shot-Prompting
 - Generierung von Ausgaben ohne existierende Beispiele
- Chain-of-Thought-Prompting
 - Verbesserung der Ausgabe durch die Einführung schrittweiser Argumentation

```
def generatePrompt(filename):  
    prompt = f'''Betrachte ausschließlich den Inhalt der Datei mit dem exakten Dateinamen {filename}  
    Als ersten Schritt soll der Inhalt der Datei {filename} betrachtet werden.  
    Basierend auf diesem Inhalt sollen Informationen ausgelesen werden. Dateien, die nicht den Datei  
  
    Diese Informationen sollen in einem JSON-Format ausgegeben werden.  
    Das JSON-Format darf keine weiteren Informationen enthalten.  
  
    Regeln, an denen sich die Befüllung der Struktur gehalten werden muss:  
    1. Die Absender und Empfänger sind Personen, die wie folgt angegeben werden sollen: <Name> (<Titel>  
    2. Existiert kein Dokortitel, dann wird nur der Name des Absenders und Empfängers verwendet.  
    4. Wenn ein Vorschlag zur Lösung des Problems angegeben wurde, notiere dies im Feld Lösungsvorschlag  
    5. Wenn kein Lösungsvorschlag angegeben wurde, lass dieses Feld leer.
```

Abbildung: Auszug aus einer beispielhaften Generierung eines Prompts zur Extraktion relevanter Informationen aus einer Datei.

Ergebnis:

- ❖ Ausgabe in einem strukturierten Format
- ❖ Informationen zu jeder Zwischenverfügung:
 - Absender, Empfänger
 - Probleme
 - Lösungsvorschläge
 - Kategorie des Problems
 - Frist zur Behebung des Problems

Aktuell:

- ❖ Kategorisierung (Welche soll es geben?)
- ❖ Weitere Analyse und Verfeinerung der Ausgabe auf verschiedenen Zwischenverfügungen (ca. 180.000)
- ❖ Verknüpfung der Ursprungsdokumente mit den extrahierten Problemen

```
{
  "Dateiname": "██████.txt",
  "Absender": "██████, Richter am Amtsgericht",
  "Empfänger": "Notar Dr. ██████",
  "Probleme": [
    {
      "Problem": "Gemäß § 40 II GmbHG ist die neue Gesellschafterliste vom  
Notar zu unterzeichnen und außerdem mit einer Notarbescheinigung gem.  
§ 40 II 2 GmbHG zu versehen.",
      "Problemkategorie": "Gesetzliche Vorschriften",
      "Lösungsvorschlag": "Vorlage der neuen Gesellschafterliste und  
Notarbescheinigung"
    }
  ],
  "Frist": "1.4.2009"
}
```

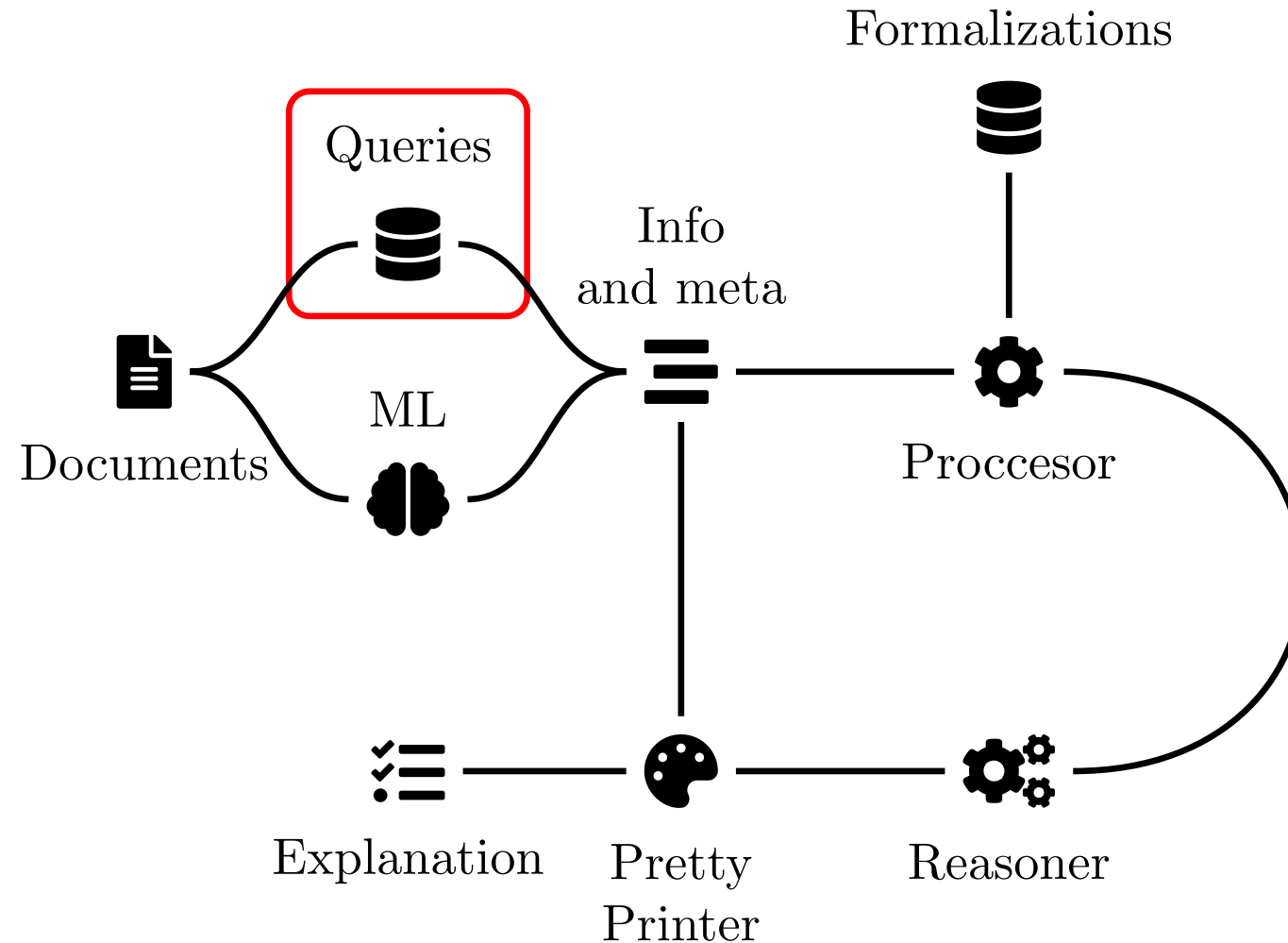
Abbildung: Ausgabe des LLM nach der Anfrage mit dem Prompt.

Computerlinguistische Analyse

von Zwischenverfügungen

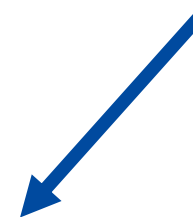
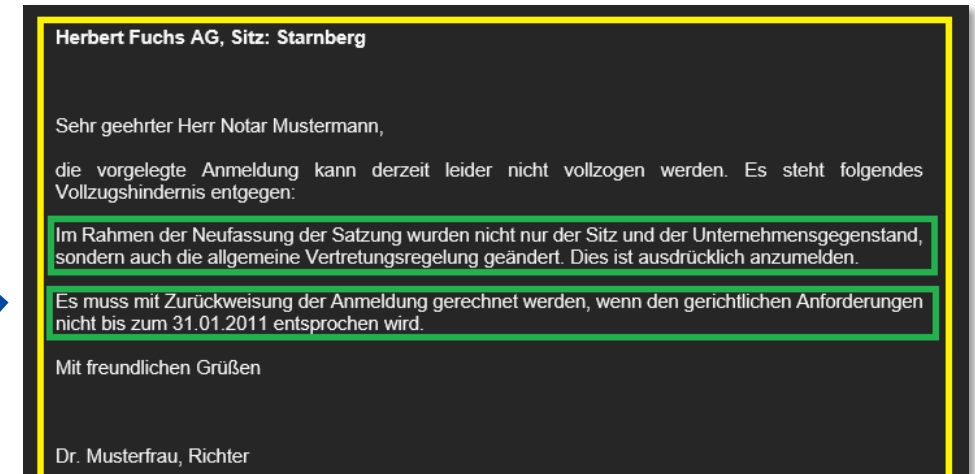
Computerlinguistische Analyse von Zwischenverfügungen

Pipeline des Registerassistenten: CCL



Aufbereitung von **120.000 Zwischenverfügungen** als **computerlinguistisches Korpus** ...

- ❖ Extraktion der Zwischenverfügungen aus Datenbank
- ❖ Konvertierung in Plain-Text-Format (→ Primärdaten)
- ❖ Tokenisierung (Erkennung von Wortgrenzen)
- ❖ Annotation mit linguistischen Merkmalen (z.B. Wortart, Lemma)
- ❖ Named Entity Recognition (z.B. Personen, Verweise auf Rechtsnormen)
- ❖ Indexierung in spezialisierter Suchoberfläche



Sehr geehrter Herr Notar [Mustermann]_{PER}, die vorgelegte Anmeldung kann derzeit leider nicht vollzogen werden . | Es steht folgendes Vollzugshindernis entgegen : Im Rahmen der Neufassung der Satzung wurden nicht nur der Sitz und der Unternehmensgegenstand , sondern auch die allgemeine Vertretungsregelung geändert . | Dies ist ausdrücklich anzumelden . | Es muss mit Zurückweisung der Anmeldung gerechnet werden , wenn den gerichtlichen Anforderungen nicht bis zum 31. 01. 2010 entsprochen wird . | Mit freundlichen Grüßen Dr. [Rechthaber]_{PER} , Richter

Zwischenverfügungen enthalten verschiedene **Elemente**, die sich nicht nur mit generativer KI extrahieren, sondern auch **linguistisch-funktional** beschreiben lassen:

- ❖ Formale Bestandteile: u.a. Anrede, Grußformel, **Feststellung der Nichtvollziehbarkeit**
(Anmeldung kann derzeit nicht vollzogen werden; Anmeldung ist nicht vollzugsfähig)

Sehr geehrter Herr Notar [Mustermann]_{PER}, **die vorgelegte Anmeldung kann derzeit leider nicht vollzogen werden** . | Es steht folgendes Vollzugshindernis entgegen : Im Rahmen der Neufassung der Satzung wurden nicht nur der Sitz und der Unternehmensgegenstand , sondern auch die allgemeine Vertretungsregelung geändert . | Dies ist ausdrücklich anzumelden . | Es muss mit Zurückweisung der Anmeldung gerechnet werden , wenn den gerichtlichen Anforderungen nicht bis zum 31. 01. 2010 entsprochen wird . | Mit freundlichen Grüßen Dr. [Rechthaber]_{PER} , Richter

Zwischenverfügungen enthalten verschiedene **Elemente**, die sich nicht nur mit generativer KI extrahieren, sondern auch **linguistisch-funktional** beschreiben lassen:

- ❖ Formale Bestandteile: u.a. Anrede, Grußformel, **Feststellung der Nichtvollziehbarkeit**
(Anmeldung kann derzeit nicht vollzogen werden; Anmeldung ist nicht vollzugsfähig)
- ❖ Information über die **Art des Fehlers**
(der Anmeldung steht aktuell noch entgegen, dass ...; es steht folgendes Vollzugshindernis entgegen: ...;)
- ❖ Erläuternde Umstände (dies ergibt sich aus §XY Abs. Z BGB, dem zufolge ...)

Sehr geehrter Herr Notar [Mustermann]_{PER}, **die vorgelegte Anmeldung kann derzeit leider nicht vollzogen werden** . | Es **steht folgendes Vollzugshindernis entgegen** : Im Rahmen der Neufassung der Satzung wurden nicht nur der Sitz und der Unternehmensgegenstand , sondern auch die allgemeine Vertretungsregelung geändert . | Dies ist ausdrücklich anzumelden . | Es muss mit Zurückweisung der Anmeldung gerechnet werden , wenn den gerichtlichen Anforderungen nicht bis zum 31. 01. 2010 entsprochen wird . | Mit freundlichen Grüßen Dr. [Rechthaber]_{PER} , Richter

Zwischenverfügungen enthalten verschiedene **Elemente**, die sich nicht nur mit generativer KI extrahieren, sondern auch **linguistisch-funktional** beschreiben lassen:

- ❖ Formale Bestandteile: u.a. Anrede, Grußformel, **Feststellung der Nichtvollziehbarkeit**
(Anmeldung kann derzeit nicht vollzogen werden; Anmeldung ist nicht vollzugsfähig)
- ❖ Information über die **Art des Fehlers**
(der Anmeldung steht aktuell noch entgegen, dass ...; es steht folgendes Vollzugshindernis entgegen: ...;)
- ❖ Erläuternde Umstände (dies ergibt sich aus §XY Abs. Z BGB, dem zufolge ...)
- ❖ **Handlungsanweisungen** (ich bitte Sie darum, ...; um Behebung binnen X Wochen wird gebeten; es muss mit Zurückweisung der Anmeldung gerechnet werden, wenn ...; es wird um Vorlage von ... gebeten)

Sehr geehrter Herr Notar [Mustermann]_{PER}, **die vorgelegte Anmeldung kann derzeit leider nicht vollzogen werden** . | Es **steht folgendes Vollzugshindernis entgegen** : Im Rahmen der Neufassung der Satzung wurden nicht nur der Sitz und der Unternehmensgegenstand , sondern auch die allgemeine Vertretungsregelung geändert . | Dies ist ausdrücklich anzumelden . | **Es muss mit Zurückweisung der Anmeldung gerechnet werden , wenn** den gerichtlichen Anforderungen nicht bis zum 31. 01. 2010 entsprochen wird . | Mit freundlichen Grüßen Dr. [Rechthaber]_{PER} , Richter

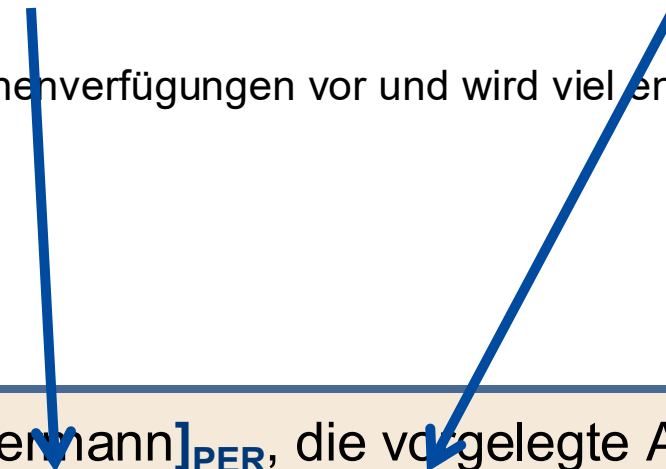
Elemente können oft mit linguistischen **Finite-State Corpus Queries** (Evert & CWB Development Team 2022) gefunden werden, z.B. **Art des Fehlers**:

Sehr geehrter Herr Notar [Mustermann]_{PER}, die vorgelegte Anmeldung kann derzeit leider nicht vollzogen werden . | Es **steht folgendes Vollzugshindernis entgegen** : Im Rahmen der Neufassung der Satzung wurden nicht nur der Sitz und der Unternehmensgegenstand , sondern auch die allgemeine Vertretungsregelung geändert . | Dies ist ausdrücklich anzumelden . | Es muss mit Zurückweisung der Anmeldung gerechnet werden , wenn den gerichtlichen Anforderungen nicht bis zum 31. 01. 2010 entsprochen wird . | Mit freundlichen Grüßen Dr. [Rechthaber]_{PER} , Richter

Elemente können oft mit linguistischen **Finite-State Corpus Queries** (Evert & CWB Development Team 2022) gefunden werden, z.B. **Art des Fehlers**:

```
[lemma="stehen"] [pos="ART|ADJ.|ADV"]* @[lemma="Vollzugshindernis"] "entgegen" ":"
```

kommt in ca. **51.000** von 120.000 Zwischenverfügungen vor und wird viel energiesparender gefunden als mit LLMs!



Sehr geehrter Herr Notar [Mustermann]_{PER}, die vorgelegte Anmeldung kann derzeit leider nicht vollzogen werden . | Es **steht folgendes Vollzugshindernis entgegen** : Im Rahmen der Neufassung der Satzung wurden nicht nur der Sitz und der Unternehmensgegenstand , sondern auch die allgemeine Vertretungsregelung geändert . | Dies ist ausdrücklich anzumelden . | Es muss mit Zurückweisung der Anmeldung gerechnet werden , wenn den gerichtlichen Anforderungen nicht bis zum 31. 01. 2010 entsprochen wird . | Mit freundlichen Grüßen Dr. [Rechthaber]_{PER} , Richter

Elemente können oft mit linguistischen **Finite-State Corpus Queries** (Evert & CWB Development Team 2022) gefunden werden, z.B. **Art des Fehlers**:

```
[lemma="stehen"] [pos="ART|ADJ.|ADV"]* @[lemma]
```

kommt in ca. **51.000** von 120.000 Zwischenverfügungen

Sehr geehrter Herr Notar [Mustermann]_{PER}, die vorg
vollzogen werden . | Es **steht folgendes Vollzugshin**
der Satzung wurden nicht nur der Sitz und der Unte
allgemeine Vertretungsregelung geändert . | Dies is
Zurückweisung der Anmeldung gerechnet werden ,
bis zum 31. 01. 2010 entsprochen wird . | Mit freund

| | |
|-------|------------------------|
| 31803 | folgendes |
| 19392 | folgende |
| 138 | weiterhin folgendes |
| 117 | noch folgendes |
| 37 | weiterhin folgende |
| 36 | noch folgende |
| 30 | folgendes weiteres |
| 24 | folgende weitere |
| 17 | immer noch folgendes |
| 16 | folgendes weitere |
| 9 | folgendes |
| 9 | nunmehr folgendes |
| 8 | immer noch folgende |
| 8 | weiter folgendes |
| 7 | noch immer folgendes |
| 5 | folgendes neue |
| 5 | nun folgendes |
| 5 | nunmehr noch folgendes |
| 4 | die folgenden |
| 4 | folgende neue |

tgegen" ":"

als mit LLMs!

zeit leider nicht
der Neufassung
ern auch die
Es muss mit
derungen nicht
er]_{PER} , Richter

Welche Fehlerkategorien kommen vor? Wie kann Abhilfe geschaffen werden?

Aktuelle Experimente: Entwicklung von **Queries** zur Extraktion spezifischer Aussagen, z.B. für **Handlungsanweisungen** eingeleitet durch *Es wird gebeten, ...; Ich bitte um ...; Ich bitte also darum, ...; Sie werden gebeten, ...; usw.*

```
[pos="PPER"] [lemma="werden"]? [lemma="bitten"] [pos="PPER|ADV" | word="um"]* ", "?  
a:[::] []+ </s> show a .. matchend;
```

mit Treffern in über **18.000** Zwischenverfügungen.

Ergebnisbeispiele:

- ❖ ... die Beteiligten zur Zahlung des Vorschusses/Kostenvorschusses zu veranlassen (12.000x)
- ❖ ... eine zustellungsfähige Anschrift anzugeben/anzumelden (900x)
- ❖ ... um Klärung/Klarstellung/Überprüfung/Berichtigung/Nachreichung/Ergänzung/... (> 200x)
- ❖ ... um Vorlage/Einreichung einer aktuellen Gesellschafterliste (70x)
- ❖ ... das/dieses Versehen zu entschuldigen (24x)

Nächste Schritte:

- ❖ Verbesserung der Korpusaufbereitung (insb. *named entity recognition*)
- ❖ Formulierung und Optimierung von Queries für unterschiedliche Aussagetypen
- ❖ Kategorisierung der Fehlerbeschreibungen und Handlungsanweisungen:
Kombination von Queries (stereotype Formeln) mit semantischem Clustering (span embeddings)
 - Wie lassen sich die verschiedenen Informationen kategorisieren und wie häufig kommen sie vor?
 - Welche Aspekte sind spezifisch für bestimmte Dokumente oder Fallarten; welche lassen sich auf andere Fälle übertragen?
- ❖ Queries für die präzise Extraktion relevanter Teilaussagen
 - korpuslinguistische Modellierung häufiger Fehler und Handlungsanweisungen
 - fachlicher Austausch mit der juristischen Projektseite essenziell:
Welche Aussagen sind wirklich äquivalent und wo müssen Unterscheidungen getroffen werden?
 - übrige Fälle dann mit Clustering oder generativen LLM analysieren
- ❖ Queries generieren auch Trainingsdaten für spezialisierte LLM → Precision-Recall-Tradeoff

Ausblick: effektive korpuslinguistische Analyse mit FlexiConc

- ❖ Forschungsprojekt *Reading Concordances in the 21st century*
- ❖ quantitative Algorithmen zur Strukturierung der Query-Resultate
- ❖ z.B. Kombination mit semantischem Clustering, manuelle Filter, ...

Corpus

DNOV-CWB

Analysis Tree

Query result (4031)

- sort: random
- ▶ **select:** within suspension (276)
 - sort: random
 - sort: right context
 - partition: by POS tag @ L1, sort: left context
- ▶ **select:** set [PRP\$] of **partition:** by POS tag @ L1 (211)
 - **sort:** left context

Query

"eyes" %c

Partition / Cluster

▼

Sort

by left context

✕

⊕

Make selection

▼

Bookmark

♥

| | | | | |
|----|--------|------------------------------------------------------------------|-------|------------------------------------------------------------------|
| 1 | OMF 38 | ollecting [, ' said Mrs Boffin , with her work abandoned , her | eyes | upon the fire , and her chin upon her hand] , ' do you recall |
| 2 | BH 49 | red thousand thanks , old fellow [, " says Mrs. Bagnet , her | eyes | sparkling with pleasure and her hand stretched out to him] |
| 3 | BH 22 | brought into my head , master [, " returns the woman , her | eyes | filling with tears] , " when I look down at the child lying so . |
| 4 | OT 24 | ' she cried , jumping fiercely up : her face flushed , and her | eyes | starting from her head--'I] robbed her , so I did ! She was n |
| 5 | OMF 42 | netimes I have fancied , sir [, ' said Bella , turning away her | eyes] | , ' that your great regard for Mrs Boffin is a very powerful m |
| 6 | DS 18 | , if ever there was one [, ' said Mrs Chick , drying both her | eyes | at once with great care on opposite corners of Miss Tox 's h |
| 7 | OT 7 | deed . ' Ah ! ' said Mr. Bumble , when the lady brought her | eyes | down to earth again] ; ' the only thing that can be done now |
| 8 | LD 2 | ' ' Not [, ' said Pet , apologetically and embarrassed by her | eyes] | , ' not , of course , that we are any company to you , or that |
| 9 | BH 8 | ng that her spectacles were made the less engaging by her | eyes | being what Ada called] " choking eyes " , " meaning very pr |
| 10 | BH 23 | Miss Summerson [, " said Mrs. Jellyby , sweetly casting her | eyes | for a moment on me and considering where to put the parti |
| 11 | PP 22 | tared wildly at her . ' Wretch [, ' said the lady , covering her | eyes | with her hands] , ' what do you want here ? ' ' Nothing , ma |
| 12 | HT 9 | not yours . ' ' Dear Miss Louisa [, ' said Sissy , covering her | eyes | , and sobbing yet] ; ' I came home from the school that afte |
| 13 | OMF 61 | ching her cheek , with a grave smile , as she cast down her | eyes | and pouted again] ; ' look at me . I want to speak to you . ' ' |
| 14 | LD 55 | eneral . ' ' Mr Dorrit [, ' said Mrs General , casting down her | eyes | as she rose with a bend] , ' must ever claim my homage an |
| 15 | DS 58 | . ' I need not remind you [, ' said Harriet , casting down her | eyes | upon her black dress] , ' through what means our circumsta |
| 16 | OMF 25 | ith stately resignation as she shook her head and dried her | eyes] | ; ' we have at present no stipendiary girl to do so . We have |
| 17 | ED 13 | ue . ' If we knew yesterday [, ' said Rosa , as she dried her | eyes] | , ' and we did know yesterday , and on many , many yester |
| 18 | DS 2 | always be a balm and comfort to me [, ' here she dried her | eyes] | ; ' it becomes necessary to provide a Nurse . ' ' Oh ! Ah ! ' s |
| 19 | TTC 44 | gement that may be [, " said Miss Pross , striving to dry her | eyes | and compose herself] , " I have no doubt it is best that Mrs. |
| 20 | DC 51 | a gleam of light upon me , Trot [, ' said my aunt , drying her | eyes] | , ' when I formed the resolution of being godmother to your |

"zum"%c [pos = "A"]? [lemma = "Geschäftsführer"]
[lemma = "werden"] [lemma = \$V_bestellen] ":"
<<PERSON>> ", "? "geb\.|geboren" "am"? <<DATUM>> ", "?
[lemma = \$V_wohnen] [pos = "Pr"] <<ORT>>

HRB Nr. 12345

Mustermann GmbH mit dem Sitz in Nürnberg

Hier: Geschäftsführerbestellung

Zur Eintragung in das Handelsregister wird bezüglich der oben genannten Gesellschaft angemeldet, was folgt:

Zum neuen Geschäftsführer wurde bestellt:

Herr Maximus **Mustermann**, geb. am 01.01.2000, wohnhaft in Musterstadt.

PERSON

DATUM

ORT

Er ist als Geschäftsführer - auch bei Vorhandensein von weiteren

Ein anderes computerlinguistisches Problem: **Anonymisierung der Texte**

- ❖ Handelsregisteranmeldungen (und auch Zwischenverfügungen) enthalten sensible personenbezogene Daten
- ❖ Experimente derzeit nur auf speziellen Rechnern ohne Internetzugang möglich (→ in verschlüsseltem Container)
- ❖ schränkt flexible Forschung und Anwendung von KI-Modellen erheblich ein
- ❖ bereits 120k Zwischenverfügungen → nur vollautomatische Anonymisierung praktikabel

HRB Nr. 12345

Mustermann GmbH mit dem Sitz in Nürnberg

Hier: Geschäftsführerbestellung

Zur Eintragung in das Handelsregister wird bezüglich der oben genannten Gesellschaft angemeldet, was folgt:

Zum neuen Geschäftsführer wurde bestellt:

Herr **Maximus Mustermann**, geb. am **01.01.2000**, wohnhaft in **Musterstadt**.

PERSON

DATUM

ORT

Er ist als Geschäftsführer - auch bei Vorhandensein von weiteren

Automatische Anonymisierung von Registerdokumenten

Anonymisierung als Voraussetzung für Legal Tech (z.B. KI-Modelle im HPC)

Vollautomatische **Anonymisierung von Gerichtsurteilen** bereits jetzt möglich!

- ❖ entwickelt in Forschungsprojekten **LeAK** (BayStMJ, 2020–2023) und **AnGer** (BMBF, 2023–2025)
- ❖ hochspezialisierte Modelle erreichen **> 99% Recall** für Namen, Datum, Adressen, ... (AG, Mietrecht/Verkehrsrecht)
- ❖ u.a. Testbetrieb am OLG Celle
- ❖ laufende Forschung zur effizienten Domänenanpassung (z.B. an Registerdokumente)

Hilfe

Risiko ?
Wie gefährlich die Veröffentlichung der Textstelle wäre.

☒ Niedrig ☒ Mittel ☒ Hoch

Mindestkonfidenz 0% ?
Je höher der Wert, desto sicherer ist die automatische Erkennung.

Tags ? Nur verwendete Tags anzeigen ☒

Sie können ganze Kategorien ("Tags") hier ausschließen.

☐ Formales • Aktenzeichen **k**

☐ Formales • Gericht **g**

☒ Natürliche Person • Name **n**

☒ Natürliche Person • Juristischer Funktionsträger **t**

Amtsgericht Erlangen

Az :11 C 122/20

Mozartstraße 23, 91052 Erlangen

Telefon: 09131/782-01

Telefax: 09131/782-105

Verkündet am: 23.7.2020

(Schneider), JAng.

Urkundsbeamtin d. Geschäftsst.

IM NAMEN DES VOLKES

In dem Rechtsstreit

Patrick Müller, Schillerstraße 24, 91054 Erlangen

- Klägerin zu 1) -

Patricia Müller, Schillerstraße 24, 91054 Erlangen

- Kläger zu 2) -

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung



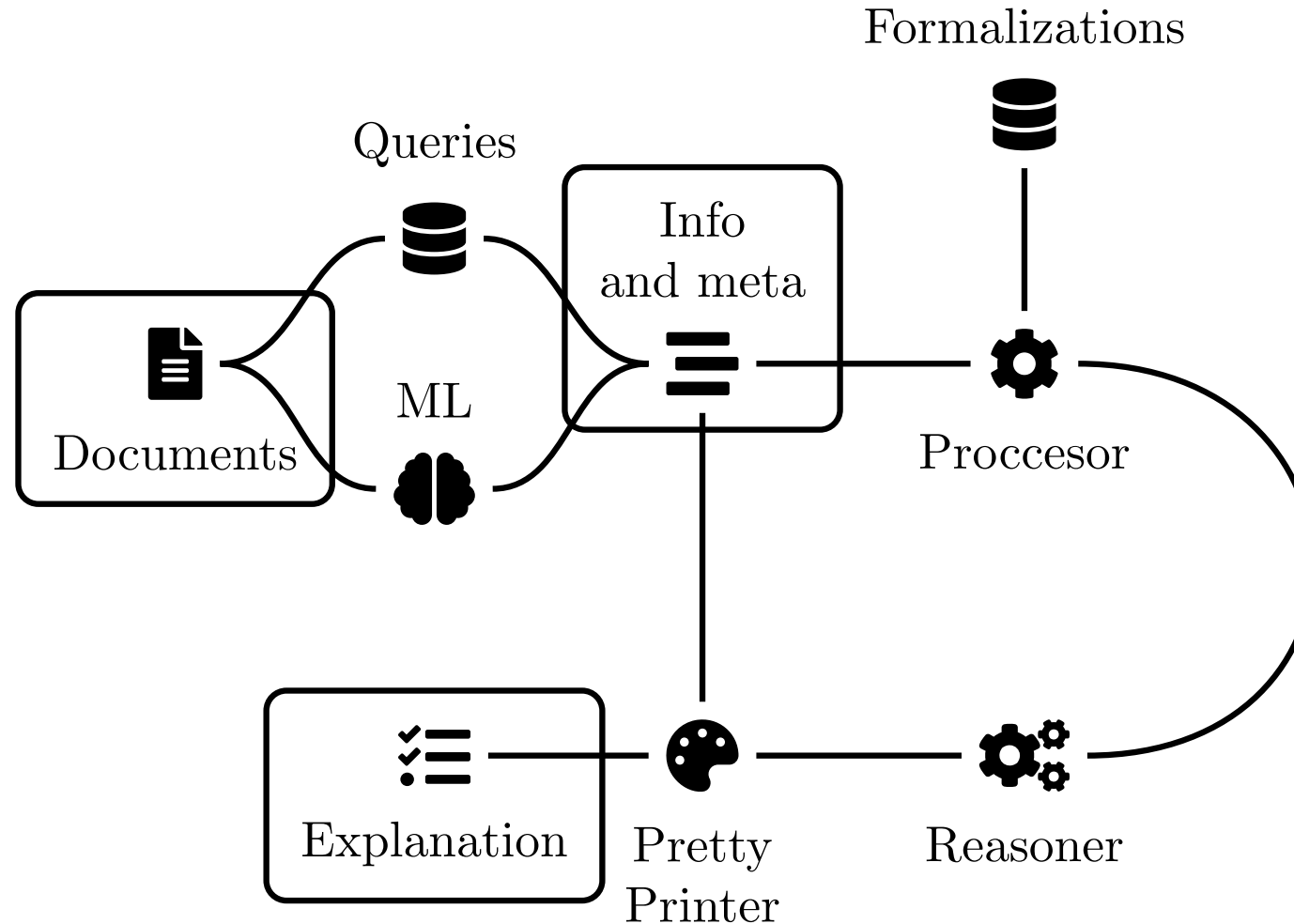
Finanziert von der
Europäischen Union
NextGenerationEU

Aufbereitung des juristischen Hintergrundwissens

Exemplarische tiefe Explizierung von Registervorgängen zur symbolischen Verarbeitung

Aufbereitung des juristischen (Hintergrund-)Wissens

Pipeline des Registerassistenten: Legal



Ziele

Ziel: Formalisierung von Registervorgängen, d.h. Übersetzung in eine formale Sprache mit dem Ziel der

- ❖ logischen Repräsentation der anzuwendenden Regeln (Phase 1) und
- ❖ automatischen Inferenz mittels dieser Regeln (Phase 2).

Beispiel zur Veranschaulichung:

Sachliche und örtliche Zuständigkeit des Gerichts

Frage 1: Ist die GmbH, hinsichtlich derer die Bestellung eines Geschäftsführers eingetragen werden soll, unter einer Registernummer des angerufenen Amtsgerichts eingetragen?

- Ja -> angerufenes Amtsgericht zuständig
- Nein -> angerufenes Amtsgericht unzuständig

```
1 #pred zustaendigkeit_des_gerichts(Gericht, Unternehmen):: 'Das Gericht @(Gericht) ist zuständig für das Unternehmen @(Unternehmen)'.
2 #pred not(zustaendigkeit_des_gerichts(Gericht, Unternehmen)):: 'Das Gericht @(Gericht) ist nicht zuständig für das Unternehmen @(Unternehmen)'.
3 #pred eintragung(Gericht, Unternehmen) :: 'Das Unternehmen @(Unternehmen) ist beim Gericht @(Gericht) eingetragen'.
4 #pred not(eintragung(Gericht, Unternehmen)) :: 'Das Unternehmen @(Unternehmen) ist nicht beim Gericht @(Gericht) eingetragen'.
5
6 zustaendigkeit_des_gerichts(Gericht, Unternehmen):- eintragung(Gericht, Unternehmen).
```

%zustaendigkeit_des_gerichts.

eintragung('Führt HRB 30456',cash_Glückspiele_Erlangen_GmbH).

Kriterien guter Formalisierung:

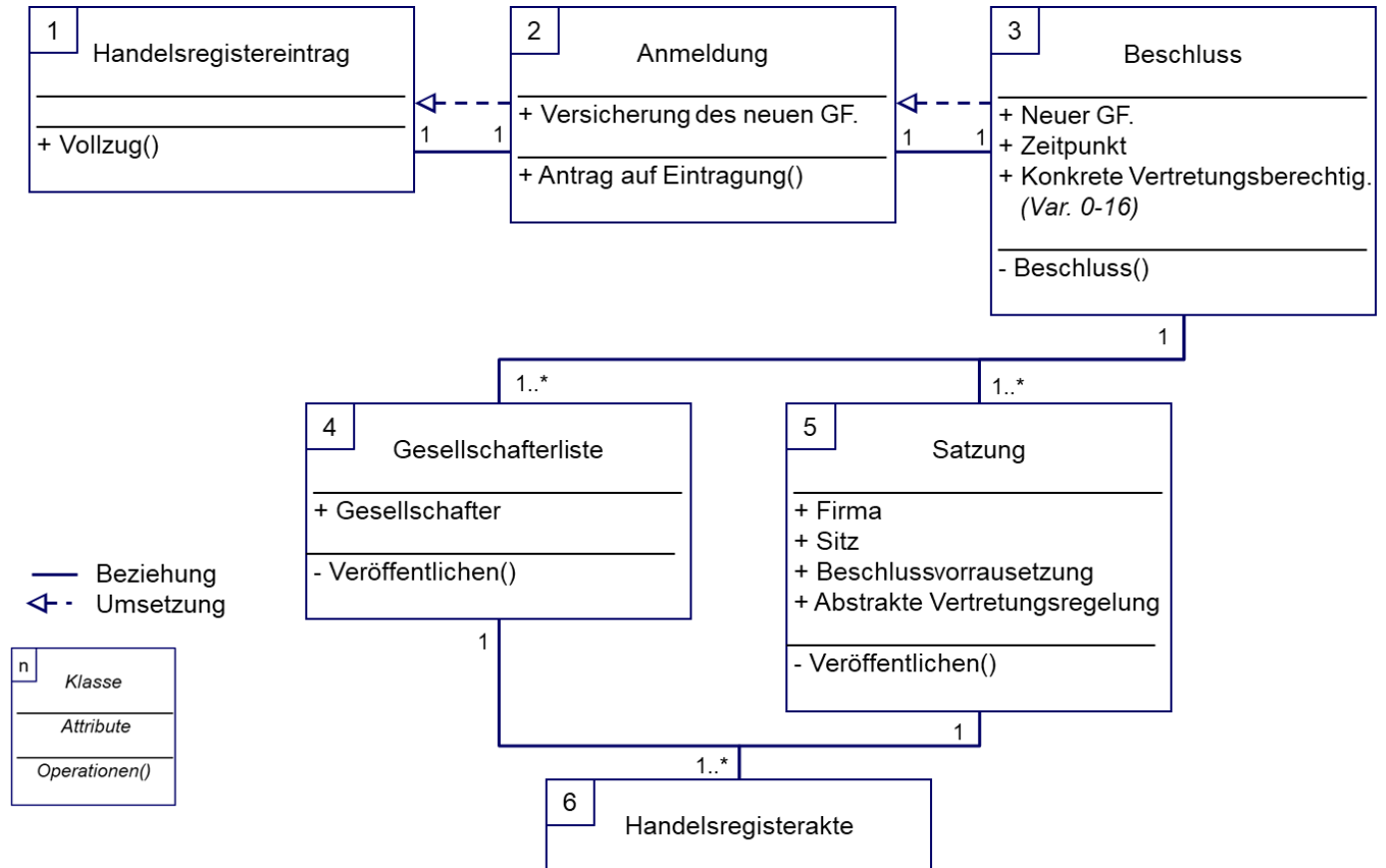
- ❖ Korrektheit
- ❖ Transparenz
- ❖ Verständlichkeit

Herausforderungen:

- ❖ Vagheit, Ambiguität und Offenheit des Rechts
- ❖ Intrinsisches und extrinsisches Material (Gesetzestext vs. Rechtsprechung vs. Registerpraxis vs. Literatur)
- ❖ „Explizites Wissen“ vs. „Implizites Wissen“
- ❖ „Soft Cases“ vs. „Hard Cases“

Erster ausgewählter Registervorgang: Anmeldung der Bestellung des Geschäftsführers einer GmbH

- ❖ überschaubarer Prüfungsumfang
- ❖ hinreichende Komplexität
- ❖ hohe Praxisrelevanz
- ❖ ausreichende Varianz
- ❖ Verbindung mit anschließend aufzubereitenden Registervorgängen (insb. Gründung der GmbH); „Concept of Basic Aspects & Atoms“ (z.B. Gesellschafterliste, -beschluss, Firma)



Aufbereitung dieser Vorgänge durch:

- ❖ Sichtung und Erfassung des einschlägigen Gesetzestextes
- ❖ Erfassung im Kontext von intrinsischem und extrinsischem Material zum Gesetzestext
- ❖ Explizierung impliziten Wissens
- ❖ Erarbeitung von Entscheidungsbäumen
- ❖ Festlegung von sog. Default-Values (Grundannahmen, die als wahr zu unterstellen sind)
- ❖ Formulierungen von Pathologien anhand von Zwischenverfügungen
- ❖ Visualisierung von Zusammenhängen

Aufbereitung des juristischen (Hintergrund-)Wissens



Schematische Aufbereitung

II.
Anspruch auf Eintragung/Eintragungspflicht

Die Bestellung des Geschäftsführers ist in das Handelsregister einzutragen, wenn die formellen und materiellen Eintragungsvoraussetzungen vorliegen.

1. Formelle Eintragungsvoraussetzungen

Die formellen Eintragungsvoraussetzungen betreffen die Registerrangmeldung und ihre Anlagen als solche. Es ist zu prüfen, ob diese formell ordnungsgemäß ist. Zu prüfen sind folgende Punkte:

a) Sachliche und örtliche Zuständigkeit des Gerichts

| Rechtlich relevante Maßstäbe (mit Normzitat): | (Teil-)Entscheidungsbaum | Beispielsfall: |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Nach § 23a Abs. 2 Nr. 3 GVG sachlich zuständig ist das Amtsgericht. | <p>Frage 1: Ist die GmbH, hinsichtlich derer die Bestellung eines Geschäftsführers eingetragen werden soll, unter einer Registernummer des angerufenen Amtsgerichts eingetragen?</p> <p>Antwort:</p> <ul style="list-style-type: none">Ja -> angerufenes Amtsgericht zuständigNein -> angerufenes Amtsgericht unzuständig | <p>Zu Frage 1:</p> <p>Die *** GmbH ist beim Amtsgericht Nürnberg unter HRB *** eingetragen.</p> <p>Die Anmeldung richtet sich an das Amtsgericht Nürnberg.</p> <p>Antwort: Ja. Amtsgericht Nürnberg ist daher zuständig.</p> |
| § 376 Abs. 1 FamFG konkretisiert dies wiederum: Zuständig ist grundsätzlich das Amtsgericht, in dessen Bezirk (auch) ein Landgericht seinen Sitz hat. Die Zuständigkeit folgt dann für den (ganzen Bezirk) dieses Landgerichts. | | |
| Eine Ausnahme zu § 376 Abs. 1 FamFG regelt wiederum § 9 GVJG: Danach ist ausnahmsweise zuständig: | | |
| <ul style="list-style-type: none">Das Amtsgericht München für die Amtsgerichtsbezirke Erding und Freising (anstatt AG Landsbut).Das Amtsgericht Fürth für Amtsgerichtsbezirke Erlangen, Fürth und Neustadt a.d. Aisch (anstatt AG Nürnberg). | | |

"Formalisierungsaufbau" (dreigliedrig)
... orientiert sich am juristischen Prüfungsaufbau

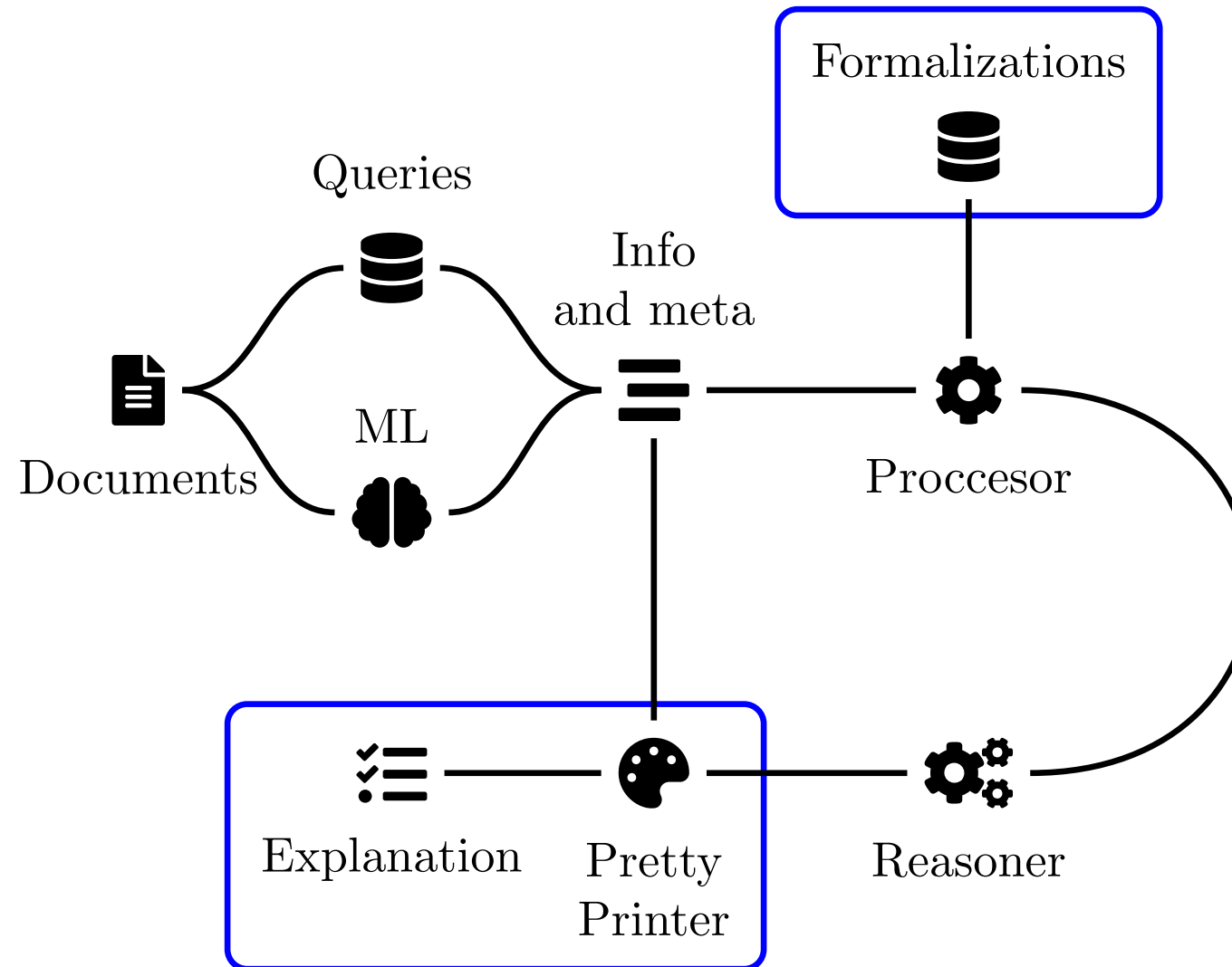
| Prüfungspunkt | zu extrahierende Informationen | Beispielsformulierungen | Normtext mit Erläuterung | Ergänzendes juristisches Hintergrundwissen | Entscheidungsbaum | Beispiel |
|------------------------------------------|--------------------------------------------|----------------------------------------------|------------------------------------------------|--------------------------------------------------|----------------------------------------------------|------------------------------------------------------|
| 1. Sachliche Zuständigkeit des Gerichts | 1.1. Sachliche Zuständigkeit des Gerichts | 1.1.1. Sachliche Zuständigkeit des Gerichts | 1.1.1.1. Sachliche Zuständigkeit des Gerichts | 1.1.1.1.1. Sachliche Zuständigkeit des Gerichts | 1.1.1.1.1.1. Sachliche Zuständigkeit des Gerichts | 1.1.1.1.1.1.1. Sachliche Zuständigkeit des Gerichts |
| 2. Örtliche Zuständigkeit des Gerichts | 2.1. Örtliche Zuständigkeit des Gerichts | 2.1.1. Örtliche Zuständigkeit des Gerichts | 2.1.1.1. Örtliche Zuständigkeit des Gerichts | 2.1.1.1.1. Örtliche Zuständigkeit des Gerichts | 2.1.1.1.1.1. Örtliche Zuständigkeit des Gerichts | 2.1.1.1.1.1.1. Örtliche Zuständigkeit des Gerichts |
| 3. Formelle Eintragungsvoraussetzungen | 3.1. Formelle Eintragungsvoraussetzungen | 3.1.1. Formelle Eintragungsvoraussetzungen | 3.1.1.1. Formelle Eintragungsvoraussetzungen | 3.1.1.1.1. Formelle Eintragungsvoraussetzungen | 3.1.1.1.1.1. Formelle Eintragungsvoraussetzungen | 3.1.1.1.1.1.1. Formelle Eintragungsvoraussetzungen |
| 4. Materielle Eintragungsvoraussetzungen | 4.1. Materielle Eintragungsvoraussetzungen | 4.1.1. Materielle Eintragungsvoraussetzungen | 4.1.1.1. Materielle Eintragungsvoraussetzungen | 4.1.1.1.1. Materielle Eintragungsvoraussetzungen | 4.1.1.1.1.1. Materielle Eintragungsvoraussetzungen | 4.1.1.1.1.1.1. Materielle Eintragungsvoraussetzungen |
| 5. Eintragungspflicht | 5.1. Eintragungspflicht | 5.1.1. Eintragungspflicht | 5.1.1.1. Eintragungspflicht | 5.1.1.1.1. Eintragungspflicht | 5.1.1.1.1.1. Eintragungspflicht | 5.1.1.1.1.1.1. Eintragungspflicht |
| 6. Eintragung | 6.1. Eintragung | 6.1.1. Eintragung | 6.1.1.1. Eintragung | 6.1.1.1.1. Eintragung | 6.1.1.1.1.1. Eintragung | 6.1.1.1.1.1.1. Eintragung |
| 7. Eintragungserfolg | 7.1. Eintragungserfolg | 7.1.1. Eintragungserfolg | 7.1.1.1. Eintragungserfolg | 7.1.1.1.1. Eintragungserfolg | 7.1.1.1.1.1. Eintragungserfolg | 7.1.1.1.1.1.1. Eintragungserfolg |

"Annotationsaufbau" (sieben-plus-n-gliedrig)
... orientiert sich am Informationsfluss der eingereichten Dokumente

- (1) Prüfungspunkt, (2) zu extrahierende Informationen,
- (3) Beispielsformulierungen, (4) Normtext mit Erläuterung,
- (5) Ergänzendes juristisches Hintergrundwissen,
- (6) Entscheidungsbaum, (7) Beispiel

Der Formalisierungsprozess

Wissensrepräsentation, Phasen und Evaluation verschiedener Tools



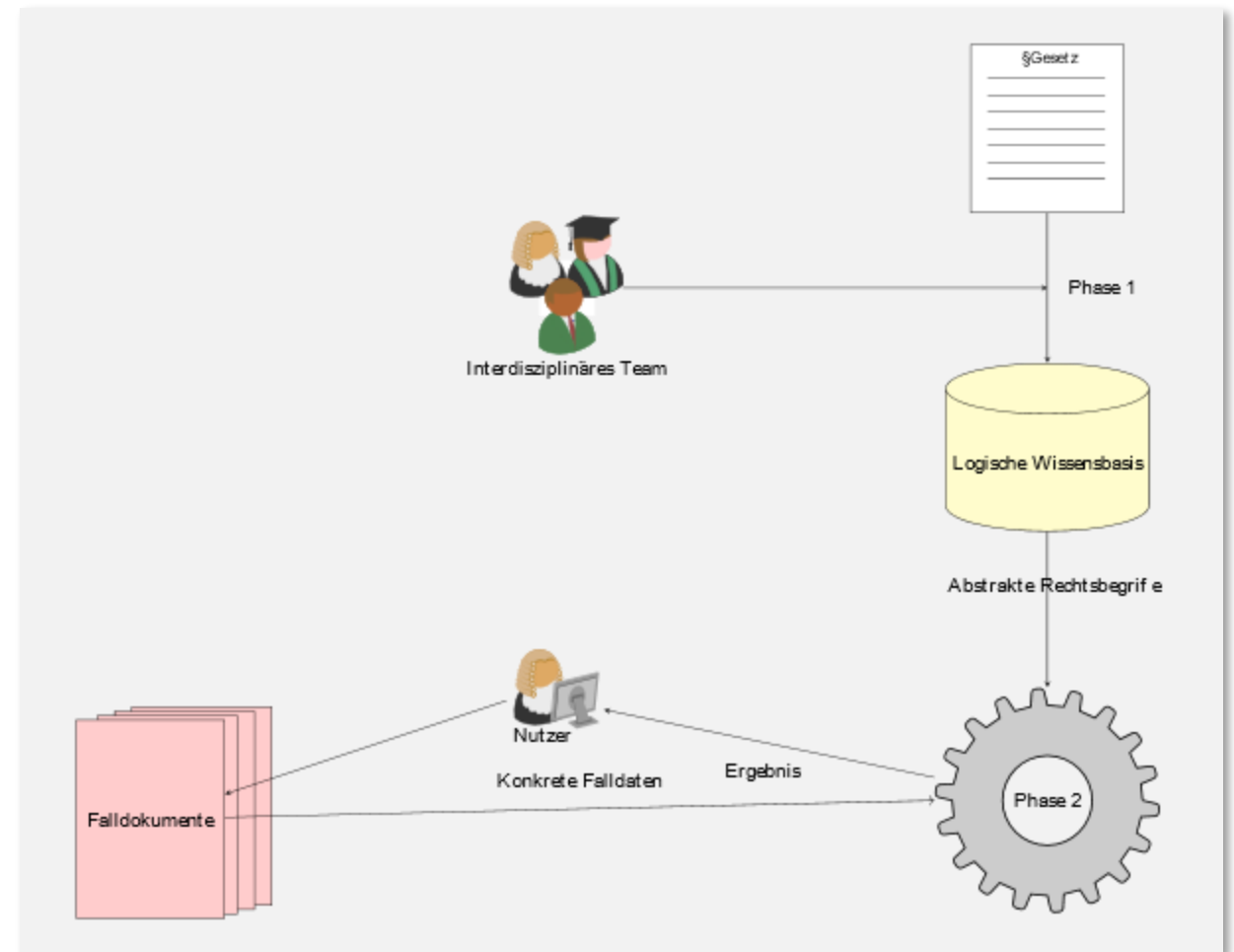


Phasen der Formalisierung nach Novotna und Libal 2022

- ❖ Phase 1: Übersetzung des Gesetzestextes in eine logische Repräsentation unter Wahrung des rechtlichen Sinngehalts
- ❖ Phase 2: Automatische (logische) Inferenz auf Basis der Ergebnisse aus Phase 1

Herausforderungen

- ❖ Phase 1: Treue und Autorität der Formalisierung
- ❖ Phase 2: Subsumtionsproblem. Inferenz nur möglich wenn konkrete, natürlich-sprachliche Fakten den abstrakten Regeln aus Phase 1 zugeordnet werden.

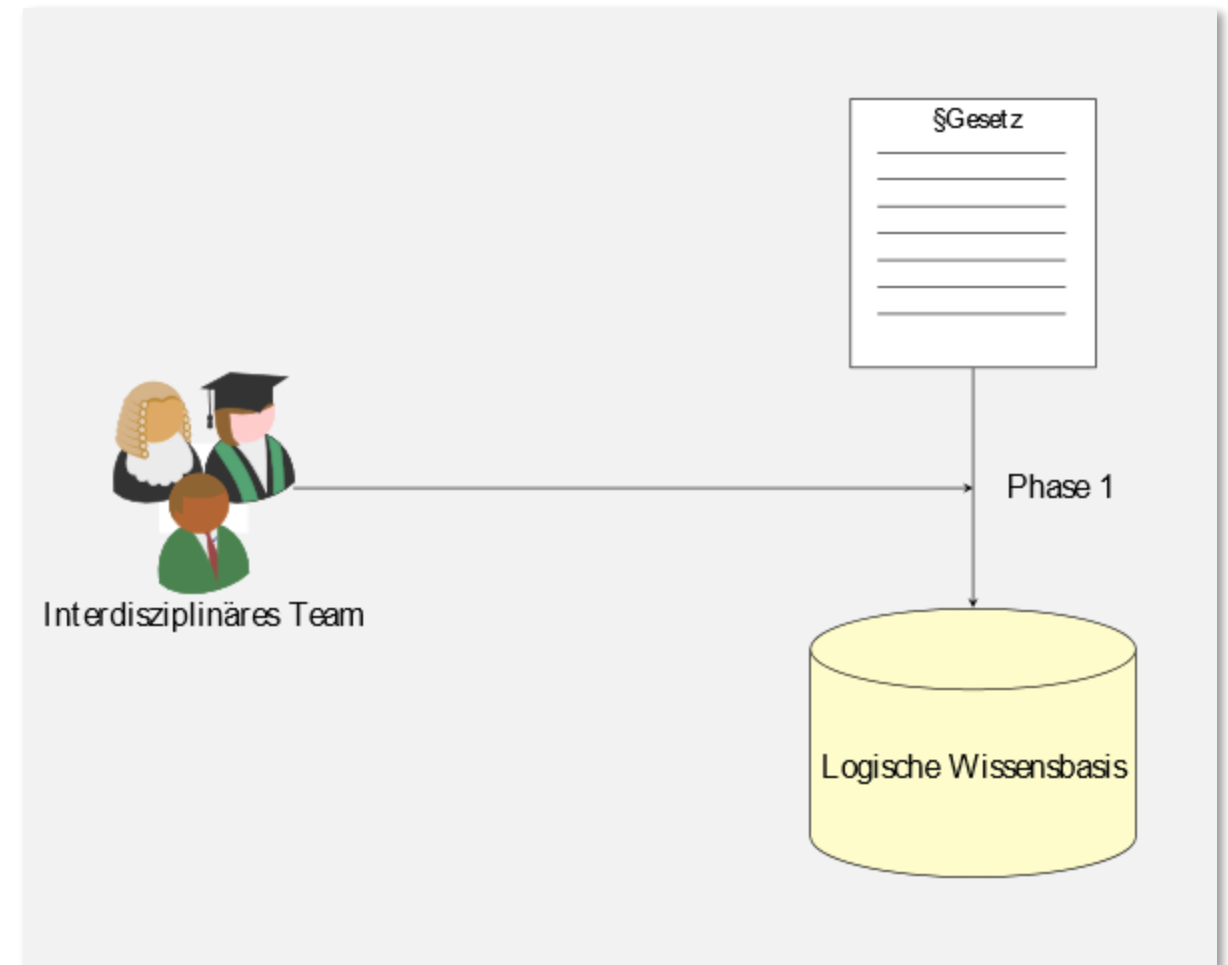


Modelle des Formalisierungsprozesses

Phase 1: Übersetzung des Gesetzestextes in eine logische Repräsentation

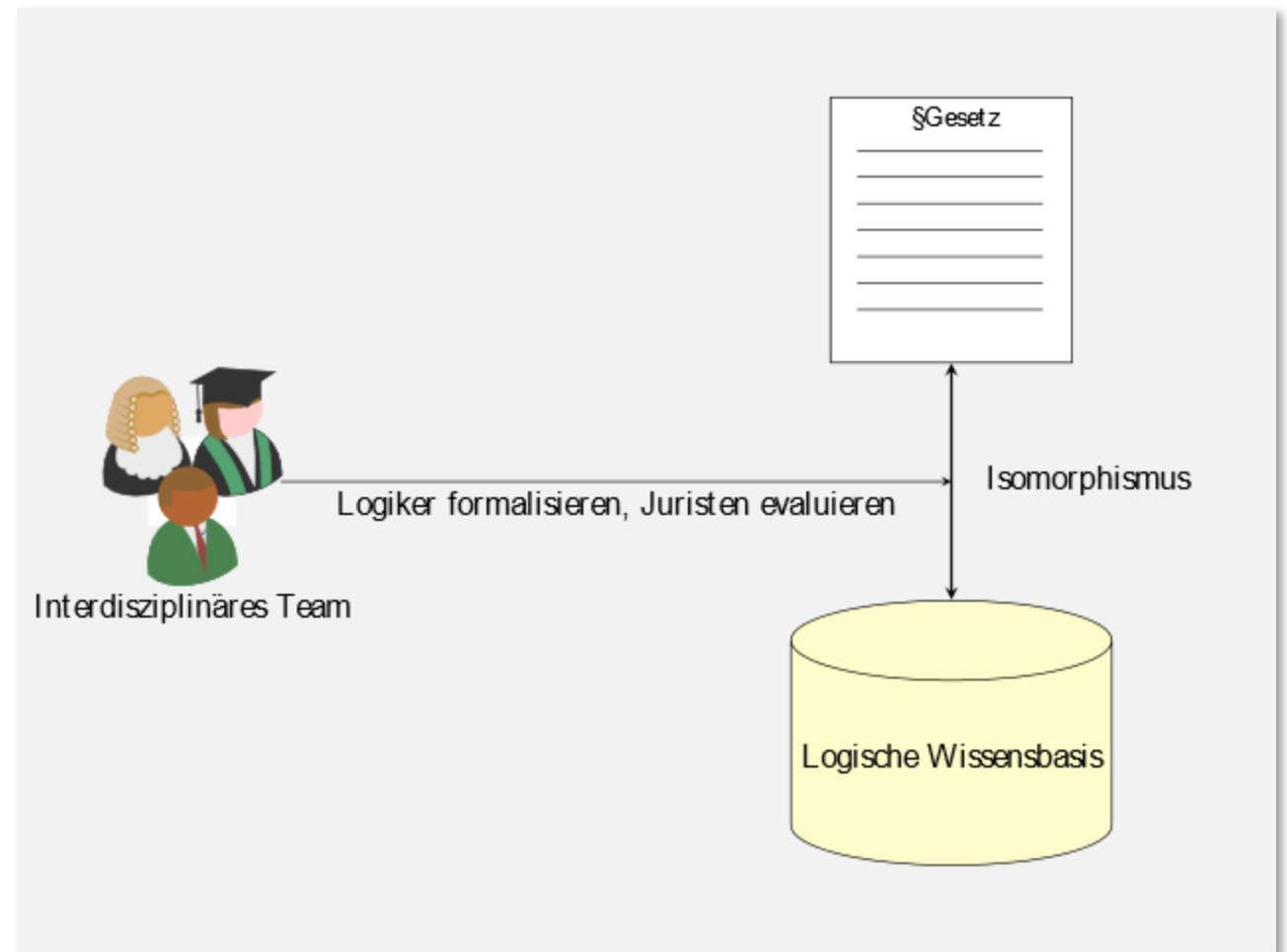
Die Literaturanalyse ergibt die folgenden **Kategorien von Modellen gelungener Übersetzungen**

- ❖ **Isomorphismusmodell:** Eins-zu-Eins- Korrespondenz zwischen Gesetzestext und seiner logischen Repräsentation
- ❖ **Interpretativ-Argumentatives Modell:** Eins-zu-n Korrespondenz im Sinne mehrere Auslegungsmethoden mit Management konfligierender Interpretationen.
- ❖ **Zwischensprachenmodell:** Eins-zu-n-zu-n' Korrespondenz in der Beziehung Gesetzestext – zwischensprachlicher Repräsentation - logischer Repräsentation wobei die Korrespondenz der Teilübersetzungeneinem einem der zuvor genannten Modelle folgen kann.



Im **Isomorphismusmodell** wird versucht möglichst nah an eine Eins-zu-Eins-Korrespondenz zwischen Gesetzestext und logischer Repräsentation zu kommen.

- ❖ Besonders geeignet wenn das System hauptsächlich mit Soft Cases umgehen muss
- ❖ Validierung durch eine interdisziplinäre Gruppe von Experten, wobei der Isomorphismus die Lesbarkeit des ``Codes'' der jeweils anderen Seite ermöglichen soll.
- ❖ Beispiel: Catala



anzuerkennen:

1. die Förderung von Wissenschaft und Forschung;
2. die Förderung der Religion;
- [...] (Aus Gründen der Übersichtlichkeit ausgelassen)
26. die Förderung der Unterhaltung und Pflege von Friedhöfen und die Förderung der Unterhaltung von Gedenkstätten für nichtbestattungspflichtige Kinder und Föten.

```
`` catala
scope SelbstlosFoerderndPruefung:
  label r52_2_S1
  # selbstlos_foerdernd :- foerderung_wissenschaft_forschung.__
  # selbstlos_foerdernd :- foerderung_religion.
  # ...
  # selbstlos_foerdernd :- foerderung_FriedhoeFe_GedenkstaetteFuerFoetenKinder.
  rule selbstlos_foerdernd under condition
    kandidat.zweck with pattern WissenschaftForschung
    or kandidat.zweck with pattern Religion
    #Aus Gründen der Übersichtlichkeit ausgelassen
    or kandidat.zweck with pattern FriedhoeFe
  consequence fulfilled
``
```

[r52_2_S2] Sofern der von der Körperschaft verfolgte Zweck nicht unter Satz 1 fällt, aber die

Beispiel: Catala (Merigoux, Chataing und Protzenko 2021)

```
[r52\_2\_S1] Unter den Voraussetzungen des Absatzes 1 sind als Förderung der Allgemeinheit
anzuerkennen:
1. die Förderung von Wissenschaft und Forschung;
2. die Förderung der Religion;
[...] (Aus Gründen der Übersichtlichkeit ausgelassen)
26. die Förderung der Unterhaltung und Pflege von Friedhöfen und die Förderung der
Unterhaltung von Gedenkstätten für nichtbestattungspflichtige Kinder und Föten.

```catala
scope SelbstlosFoerderndPruefung:
 label r52_2_S1
 # selbstlos_fördernd :- förderung_wissenschaft_forschung.__
 # selbstlos_fördernd :- förderung_religion.
 # ...
 # selbstlos_fördernd :- förderung_Friedhöfe_GedenkstätteFürFötenKinder.
 rule selbstlos_foerdernd under condition
 kandidat.zweck with pattern WissenschaftForschung
 or kandidat.zweck with pattern Religion
 #Aus Gründen der Übersichtlichkeit ausgelassen
 or kandidat.zweck with pattern Friedhoefe
 consequence fulfilled
...

[r52_2_S2] Sofern der von der Körperschaft verfolgte Zweck nicht unter Satz 1 fällt, aber die
dieser Zweck für gemeinnützig erklärt werden.
consequence not fulfilled
```

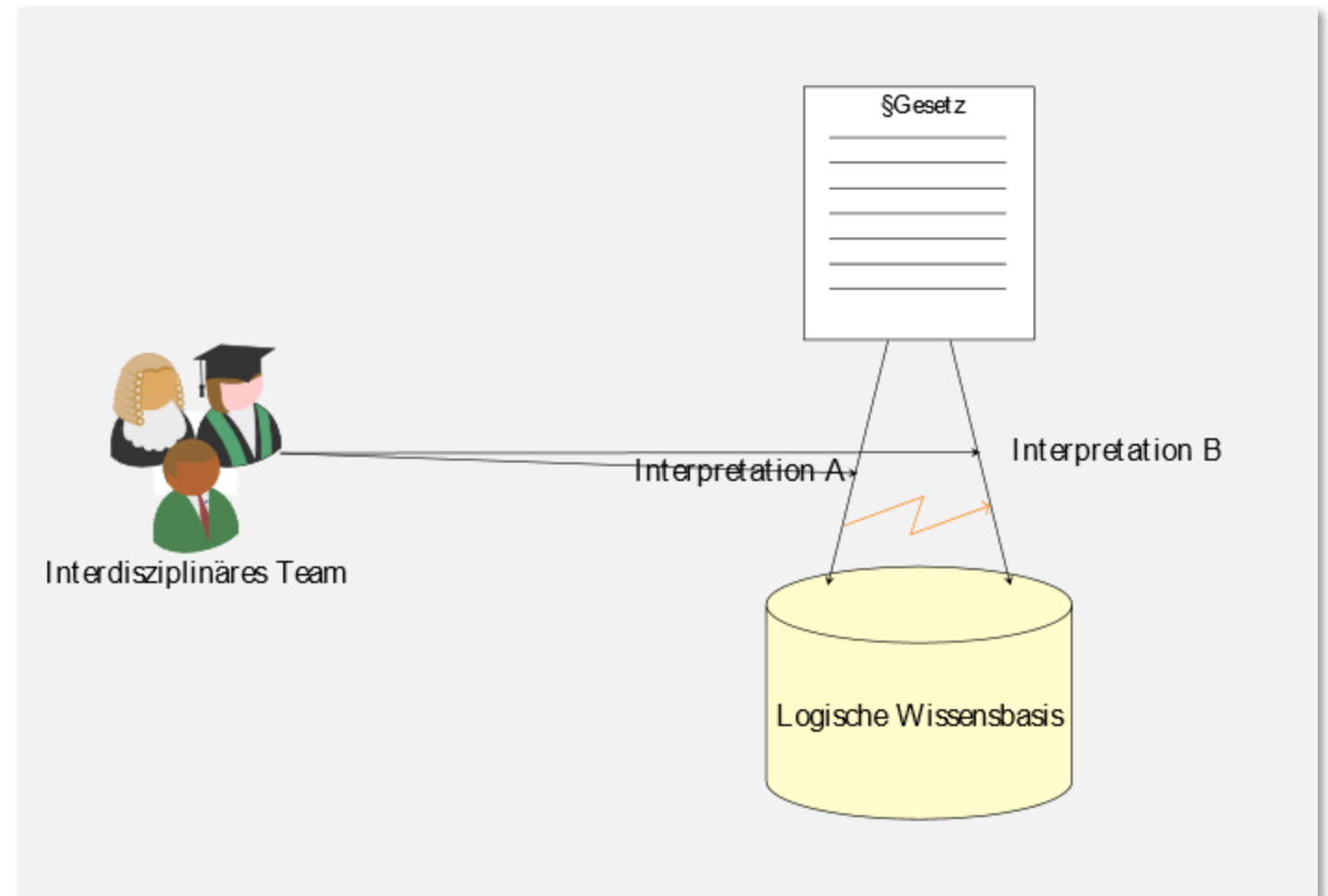
- ❖ Textnahe  
Formalisierung durch  
Literate Programming
- ❖ Ausnahmen durch  
Implementierung von  
Defaultlogik
- ❖ Backpatching via  
“context” keyword
- ❖ “Out-of-Order”  
Definitionen

Denis Merigoux, Nicolas Chataing, und Jonathan Protzenko.  
2021. Catala: a programming language for the law. In: Proc.  
ACM Program. Lang. 5, ICFP, Artikel 77 (August 2021), 29S  
Seiten.

## Zur Phase 1: Interpretativ-Argumentatives Modell

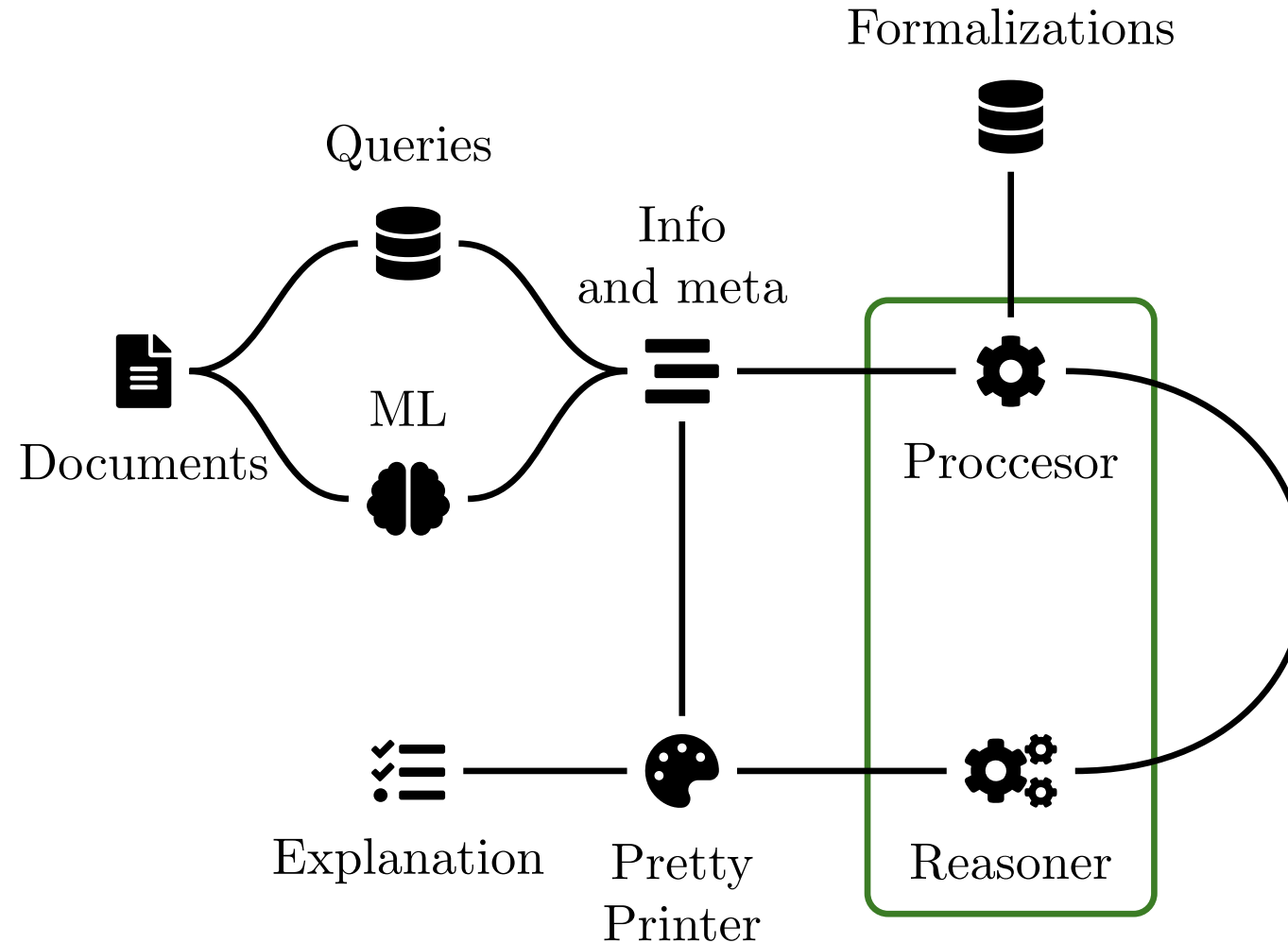
Im **Interpretativ-Argumentativen Modell** geht es darum, verschiedene, konfligierende Interpretationen desselben Textes zu modellieren.

- ❖ Für **Phase 2**: Identifizierung *einer* zumindest im gegenwärtigen Kontext korrekten Interpretation (Ausnahme: alle möglichen Interpretationen liefern dasselbe Resultat)
- ❖ **Argumentationsprozess**: Unterstützung von und Angriffe auf unterschiedliche Interpretationen basierend auf Auslegungsmethodik (i.e. Canones).
- ❖ Am besten geeignet, wenn die Unterstützung von **Hard Cases** notwendig ist (aber: siehe auch Subsumtionsproblem!).
- ❖ Beispiel: Argumentationsschemata / Carneades



# Modelle des Formalisierungsprozesses

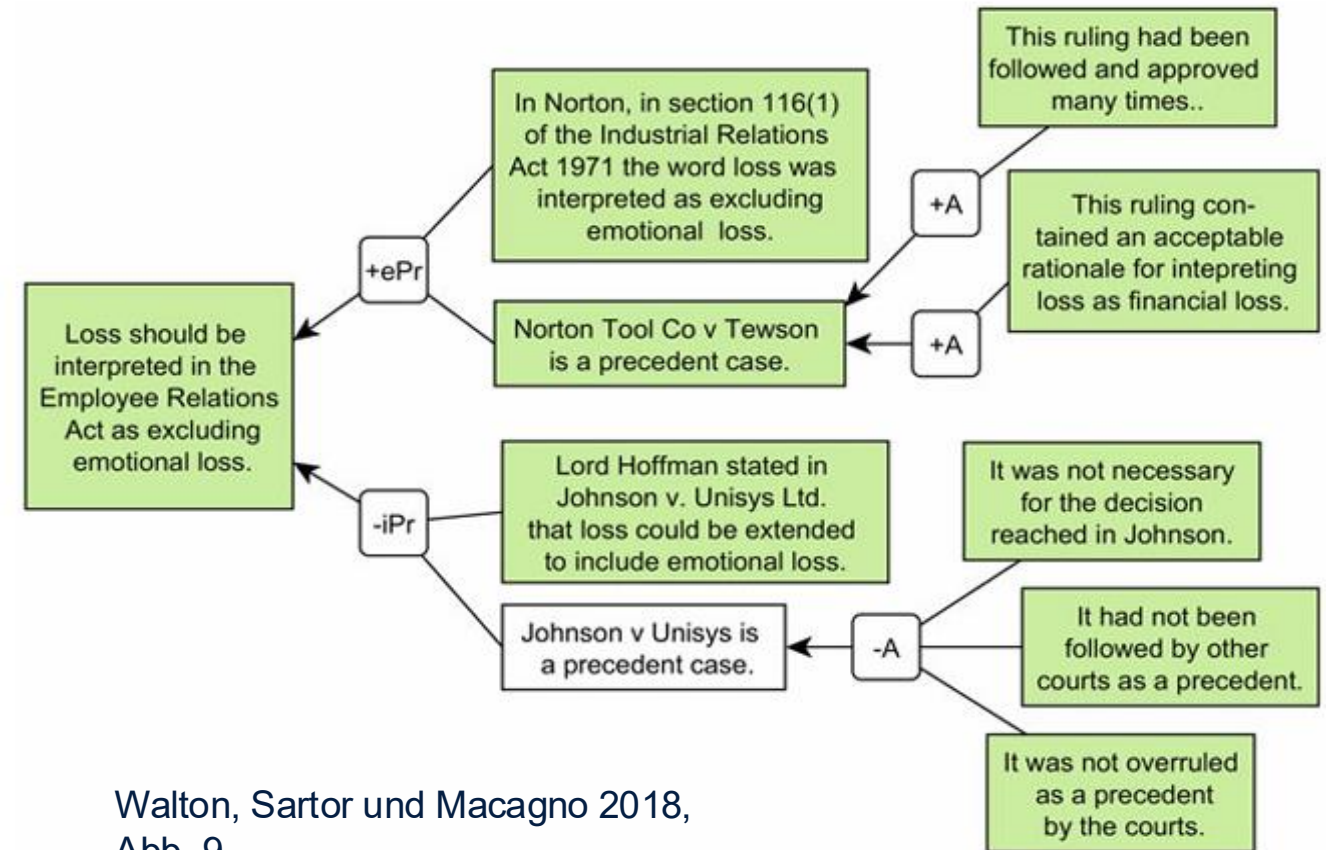
Pipeline des Registerassistenten: T.SC



# Beispiel: Carneades

Walton, Sartor und Macagno 2018

- ❖ Formalisierung Juristischer **Argumentationsschemata** als anfechtbare Schlussregeln.
- ❖ Obersätze: **Canones**.
- ❖ Untersätze: gegebene Auslegung eines Begriffs erfüllt im Dokumentkontext Prämisse des Obersatzes.
- ❖ Konstruktion aller möglichen interpretativen Argumente durch Generierung eines **logischen Abschlusses** unter Anwendung aller Schlussregeln
- ❖ Entfernung nicht haltbarer Interpretationen durch Anwendung von Ausnahmen (**Argumentationssemantik**).

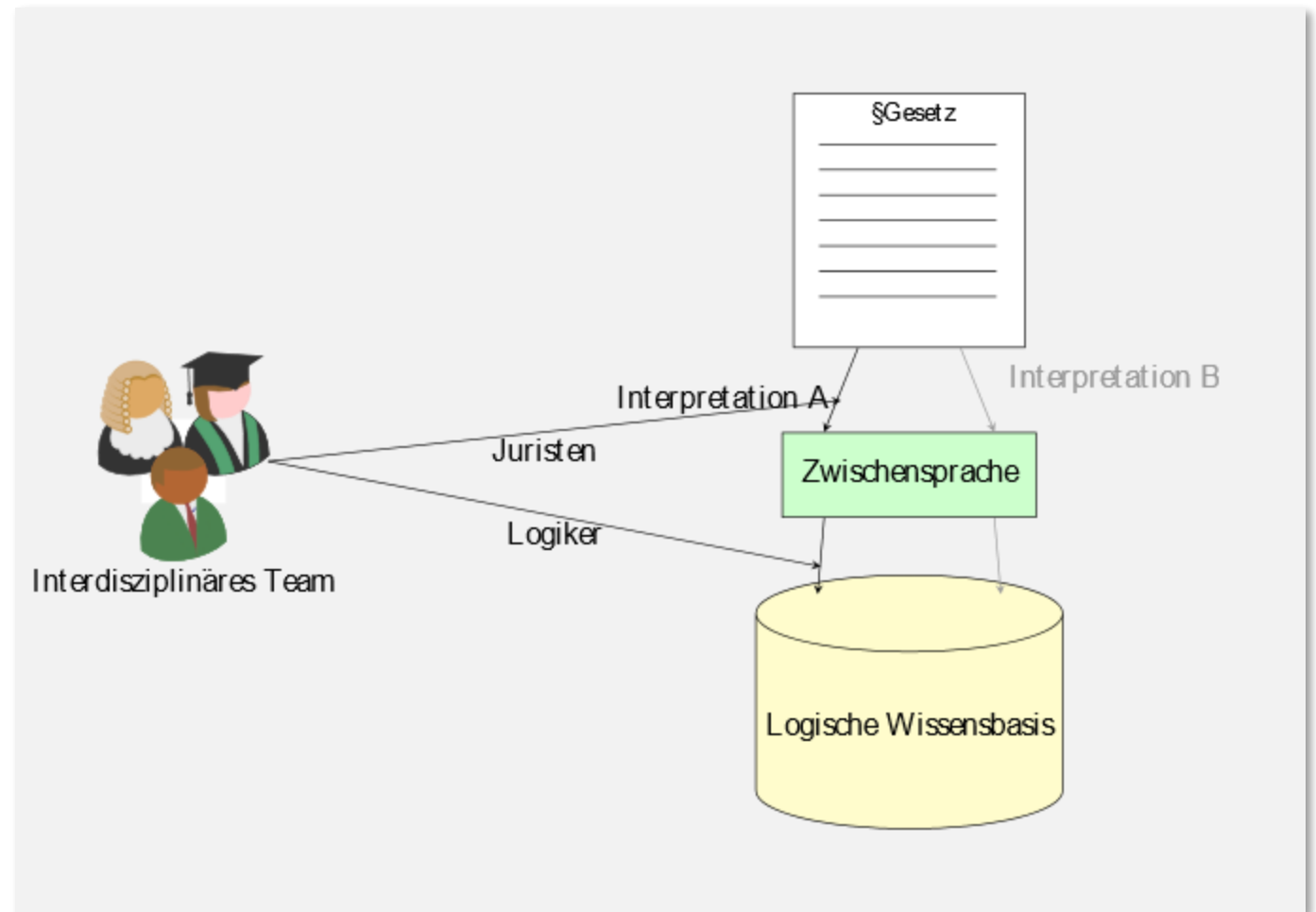


Walton, Sartor und Macagno 2018,  
Abb. 9.

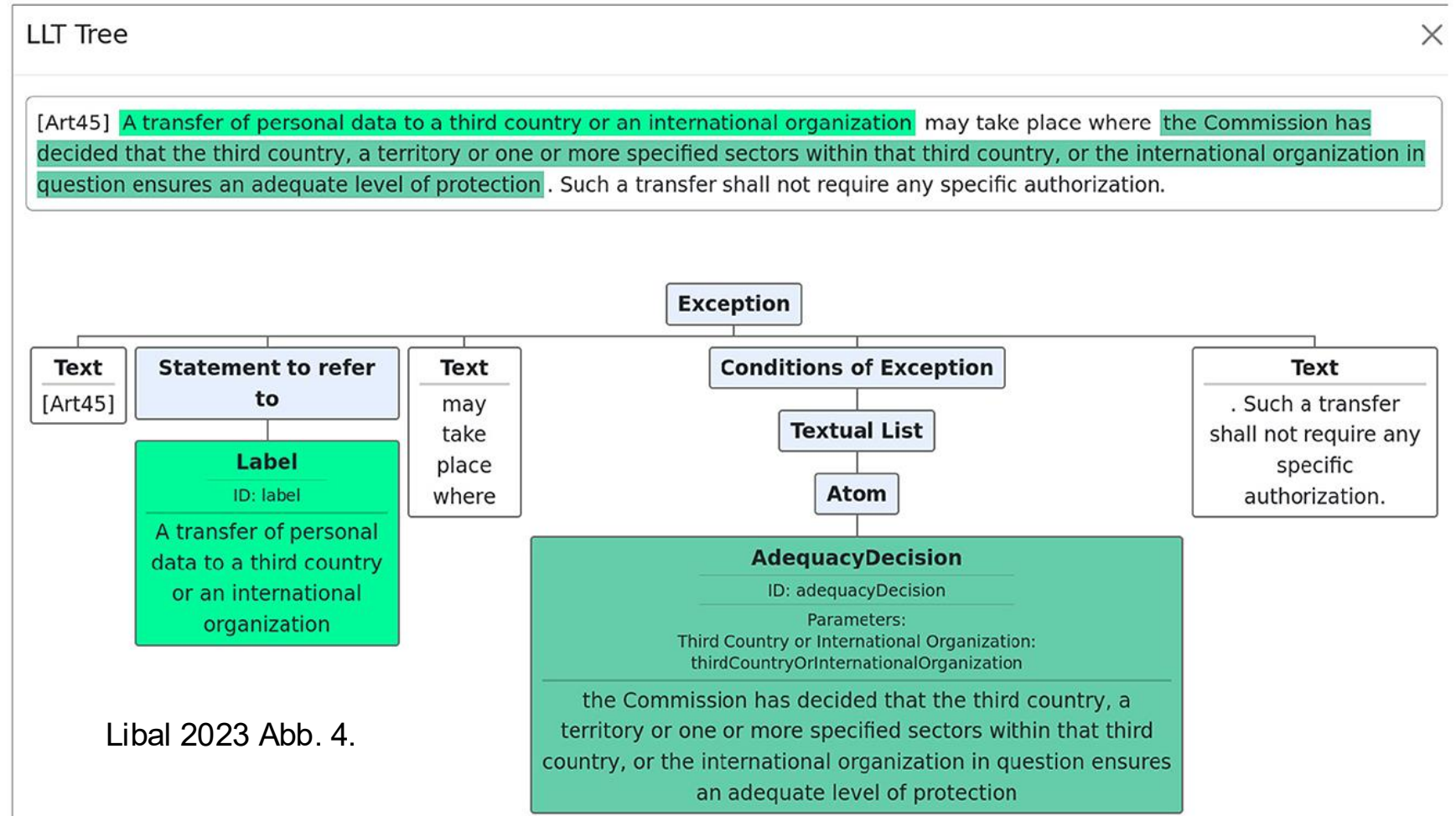


Im **Zwischensprachenmodell** ist die Hauptmotivation, das Transparenzproblem anzugehen.

- ❖ Zwischensprache für Legal und Informatik verständlich.
- ❖ Möglichkeit einer bidirektionalen Übersetzung.
- ❖ Integration des **Interpretativ-Argumentativen Modells** vor oder nach der Zwischensprache (in der Praxis meist Eins-zu-Eins-zu-Eins Modelle).
- ❖ Vorteil: Zwischensprache ermöglicht Flexibilität beim logischen Formalismus.
- ❖ **Problem:** Wahl der Zwischensprache: Kontrolliert-natürliche Sprache (CNL) / Syntactic Sugar für Logik vs. seichte Mark-Up Semantik (z.B. Argdown).



- ❖ Familie von formalen Grammatiken.
- ❖ Kann durch Hinzufügen neuer Templates erweitert werden.
- ❖ Keine fixe Semantik ...
- ❖ ... je nach Ausdruckskraft Einbettungen in u.a. Prolog oder Logik höherer Stufe möglich.
- ❖ Verschiedene Interpretationen möglich ...
- ❖ ... aber kein Konfliktmanagement.

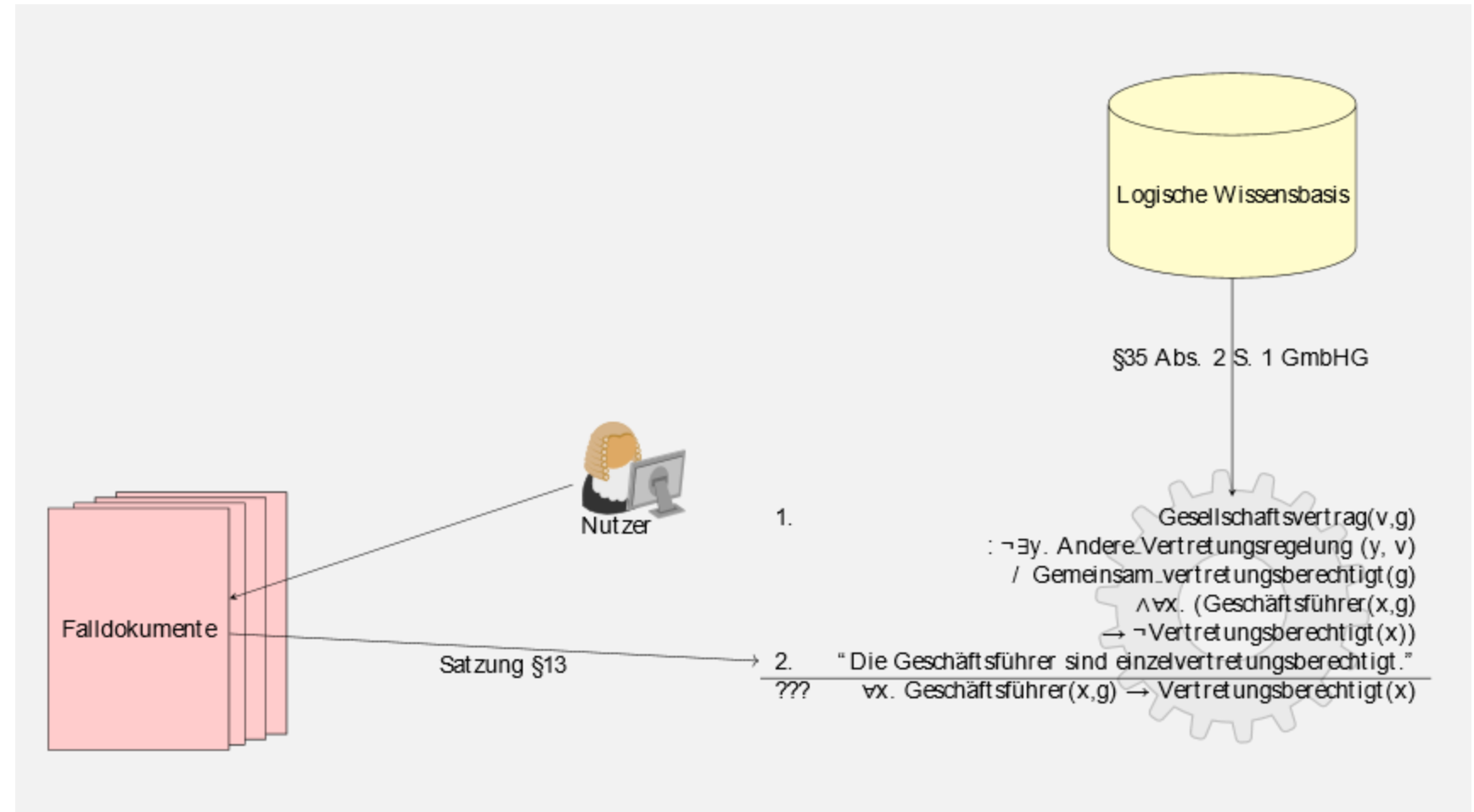




Wie können konkrete Fakten eines Falles den abstrakten, in einer Wissensbasis formalisierten Rechtsbegriffen und Normen **zugeordnet** werden?

### Ansätze

- ❖ Semantische Formulare
- ❖ NLP + Information Retrieval



### Constraint Handling Rules (CHR)

- ❖ Entwickelt als Sprache zur Implementierung von Constraint Solvern
- ❖ Kann Argumentations Schemata effizient kodieren und instanziiieren

### Constrained Answer Set Programming (CASP)

- ❖ CASP verbindet Logische Regeln mit Konkreten Bedingungen und Einschränkungen
- ❖ Produziert schöne Herleitungsbäume als Begründung

### Prolog

- ❖ Etablierte Logikprogrammiersprache basierend auf Logik erster Stufe
- ❖ Bibliotheken sowohl für CHR als auch s(CASP)
- ❖ Bietet gute Basis für Experimente

#### ▶ s(CASP) model

#### ▼ s(CASP) justification

[Expand All](#) [+1](#) [-1](#) [Collapse All](#)

▼ Die Anmeldung antrag erfüllt alle Materiellen und Formellen Voraussetzungen bezüglich der bestellung\_des\_Geschäftsführers, because

▼ Die Anmeldung antrag erfüllt alle Formellen Voraussetzungen bezüglich der bestellung\_des\_Geschäftsführers, because

▼ Das Gericht Führt HRB 30456 ist zuständig für das Unternehmen cash\_Glückspiele\_Erlangen\_GmbH, because

Das Unternehmen cash\_Glückspiele\_Erlangen\_GmbH ist beim Gericht Führt HRB 30456 eingetragen

▶ Der Antrag antrag enthält einen Verfahrensantrag, because

▶ Der Antrag antrag ist abstract eintragungsfähig für die Person johnny\_Cash, because

▼ Die Unterlagen des Antrags antrag enthalten alle notwendigen Angaben und die Erklärung um johnny\_Cash zum Geschäftsführer zu bestellen, because

▼ Alle notwendigen Daten der Person johnny\_Cash sind vorhanden, because

Der Name von johnny\_Cash ist gegeben, and

Der Familienname von johnny\_Cash ist gegeben, and

Das Geburtsdatum von johnny\_Cash ist gegeben, and

Die Adresse von johnny\_Cash ist gegeben, and

johnny\_Cash ist eine natürliche Person

Die Person johnny\_Cash hat die Erklärung nach Paragraph 6 Absatz 2 Satz 2 abgegeben

▶ Die Anmeldung antrag wurde von hinreichend Vertretungsberechtigten Personen des Unternehmens A Unterschrieben, because

▶ Der Antrag antrag enthält alle notwendigen Unterlagen für eine Bestellung des Geschäftsführers, because

▶ Die Anmeldung antrag erfüllt alle Materiellen Voraussetzungen bezüglich der

### Regeln

- ❖ Kodieren die Regulierungen und Gesetze unter Verwendung einer geeigneten Sprache

```
#pred pruefung_formelle_eigenschaften(Antrag,
Gericht, Unternehmen, Per, Zweck)::'Die Anmeldung
@(Antrag) erfüllt alle Formellen Voraussetzungen
bezüglich der @(Zweck)'.
#pred not(pruefung_formelle_eigenschaften(Antrag,
Gericht, Unternehmen, Per, Zweck))::'Die Anmeldung
@(Antrag) erfüllt nicht alle Formellen
Voraussetzungen bezüglich der @(Zweck)'.

pruefung_formelle_eigenschaften(Antrag, Gericht,
Unternehmen, Per, Zweck):-
 zustandigkeit_des_gerichts(Gericht,
Unternehmen), stellung_verfahrensantrag(Antrag),
abstract_eintragungsfahigkeit(Antrag, Per),
vollstaendigkeit_der_angaben(Antrag, Per, Zweck),
noetige_unterschriften(Antrag, Per, Unternehmen),
notwednige_unterlagen(Antrag, Zweck).
```

### Falldaten

- ❖ Kodieren die Information des konkreten Falles in der Sprache der Formalisierung

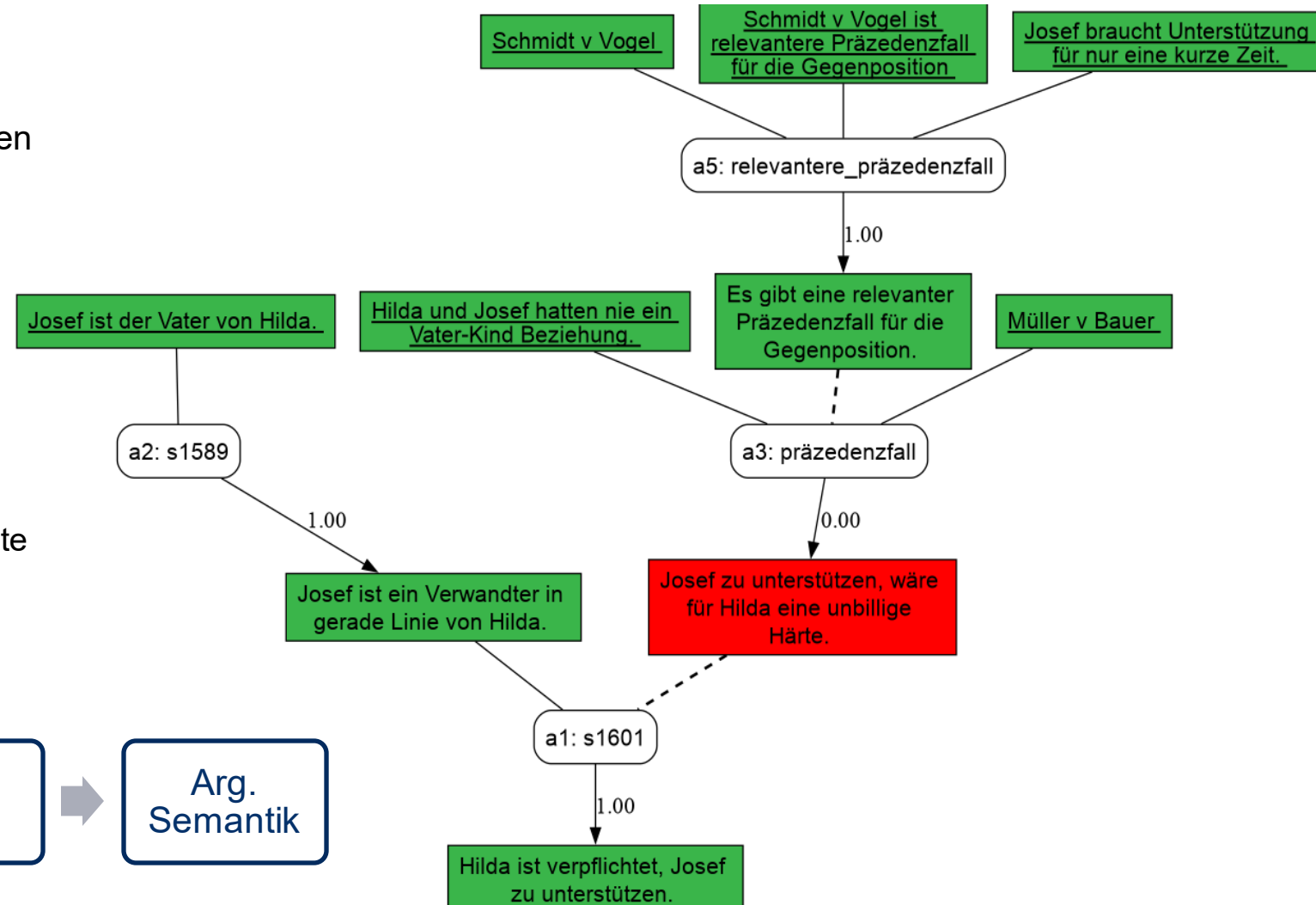
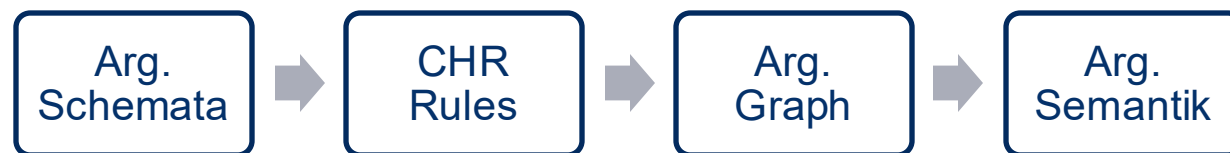
```
zugehoerige_unterlagen(beschluss,antrag).
gesellschaftsbeschluss(beschluss).

gesellschaftsliste(cash_Glückspiele_Erlangen_GmbH
, 'Alexander').
unterschrift('Alexander', beschluss).

eintragung('Führt HRB
30456', cash_Glückspiele_Erlangen_GmbH).
```

- ❖ Erste Version 2006 in einem EU Projekt zur Standardisierung juristischer Formalisierungen
- ❖ Entwicklung der aktuellen Version bis 2022
- ❖ Basiert auf CHR + formaler Argumentationssemantik
- ❖ Erweitert den Prolog+CHR Ansatz durch Visualisierung und Expressivität
- ❖ Gesetze werden in Arg. Schemata formuliert
- ❖ Falldaten werden als Annahmen formuliert
- ❖ Alternativ können auch konkrete Einzelargumente formuliert werden

### Mehrschrittiger Prozess



**Vielen Dank**  
**für Ihre Aufmerksamkeit!**

1fcfc183-5d01-490c-9e15-  
b5934a68b590