

# Vom Gesetz zur formalen Beschreibung

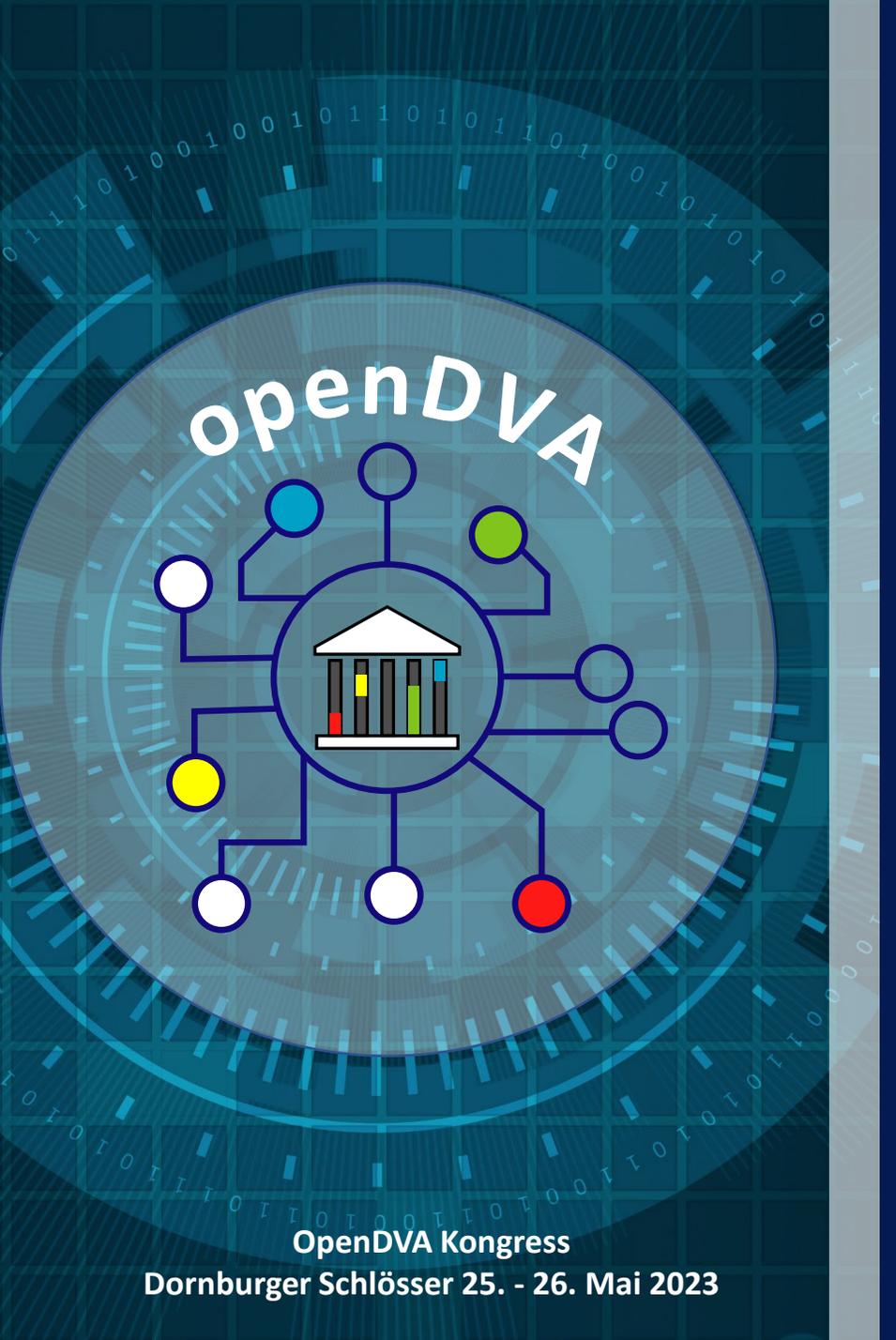
Leila Feddoul

Clara Lachenmaier

25.05.2023

OpenDVA Kongress  
Dornburger Schlösser 25. - 26. Mai 2023





# Einführung und Motivation

OpenDVA Kongress  
Dornburger Schlösser 25. - 26. Mai 2023





# Motivation

Die **Normenanalyse** ist ein wichtiger Schritt bei der Erstellung/Anpassung neuer **Verwaltungsleistungen** (z.B. Anmeldung eines Hauptwohnsitzes)

Nichtamtliches Inhaltsverzeichnis

**Gesetz zur Verhütung und Bekämpfung von Infektionskrankheiten beim Menschen (Infektionsschutzgesetz - IfSG)**

**§ 43 Belehrung, Bescheinigung des Gesundheitsamtes**

(1) Personen dürfen gewerbsmäßig die in § 42 Abs. 1 bezeichneten Tätigkeiten erstmalig nur dann ausüben und mit diesen Tätigkeiten erstmalig nur dann beschäftigt werden, wenn durch eine nicht mehr als drei Monate alte Bescheinigung des Gesundheitsamtes oder eines vom Gesundheitsamt beauftragten Arztes nachgewiesen ist, dass sie

- über die in § 42 Abs. 1 genannten Tätigkeitsverbote und über die Verpflichtungen nach den Absätzen 2, 4 und 5 vom Gesundheitsamt oder von einem durch das Gesundheitsamt beauftragten Arzt belehrt wurden und
- nach der Belehrung im Sinne der Nummer 1 in Textform erklärt haben, dass ihnen keine Tatsachen für ein Tätigkeitsverbot bei ihnen bekannt sind.

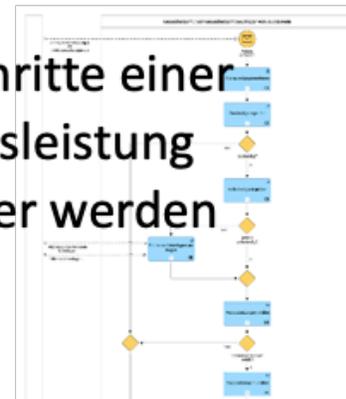
Liegen Anhaltspunkte vor, dass bei einer Person Hinderungsgründe nach § 42 Abs. 1 bestehen, so darf die Bescheinigung erst ausgestellt werden, wenn durch ein ärztliches Zeugnis nachgewiesen ist, dass Hinderungsgründe nicht oder nicht mehr bestehen.

(2) Treten bei Personen nach Aufnahme ihrer Tätigkeit Hinderungsgründe nach § 42 Abs. 1 auf, sind sie verpflichtet, dies ihrem Arbeitgeber oder Dienstherrn unverzüglich mitzuteilen.

(3) Müssen der Arbeitgeber oder Dienstherr Anhaltspunkte oder Tatsachen



- Prozessschritte einer Verwaltungsleistung  
- Datenfelder werden benötigt



Online-Verwaltungsleistungen

Gesetze, Rechtsverordnungen, etc.



# Problem

Manuelle Normenanalyse ist **zeitaufwendig** und damit **teuer**

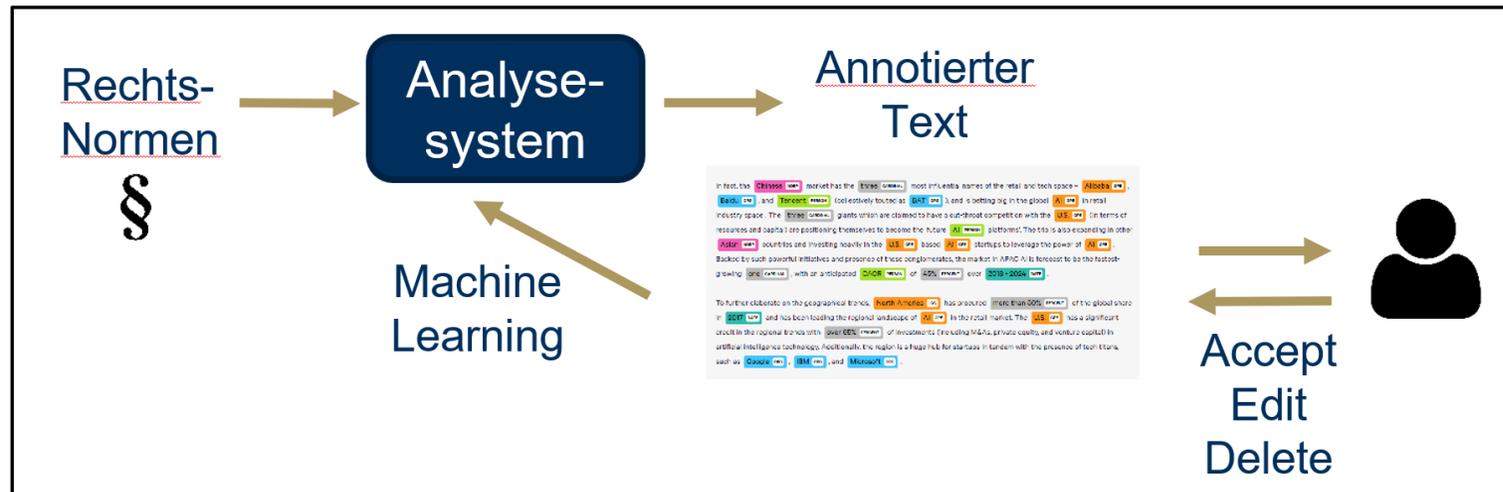
(2) Die **Träger** der **Schülerbeförderung** haben die **in ihrem Gebiet wohnenden** Schülerinnen und Schüler  
1. der allgemeinbildenden Schulen bis einschließlich des 10. Schuljahrganges; **die der Förderschulen darüber hinaus,**  
2. des **Berufsvorbereitungsjahres** und  
3. des **ersten Schuljahrganges** derjenigen Berufsfachschulen, zu deren Zugangsvoraussetzungen kein mittlerer Schulabschluss gehört, **unter zumutbaren Bedingungen** zur Schule **zu befördern** oder ihren **Erziehungsberechtigten** die notwendigen Aufwendungen für den Schulweg **zu erstatten**.  
Die Beförderungs- oder Erstattungspflicht besteht nur für die **Wegstrecke zwischen der Wohnung** der **Schülerin oder des Schülers** und der nächstgelegenen Schule der von ihr oder ihm gewählten Schulform.  
Bei der Ermittlung der nächstgelegenen Schule werden **Schulen in freier Trägerschaft** dann nicht berücksichtigt, wenn die **Schülerin oder der Schüler** eine öffentliche Schule besucht.  
Besucht die **Schülerin oder der Schüler** eine **Schule mit inhaltlichen Schwerpunkten** gemäß § 5 Abs. 1 Satz 3 oder § 6 Abs. 1 Satz 3, besteht die Beförderungs- oder Erstattungspflicht bis zur nächstgelegenen Schule mit diesem Bildungsangebot.  
Als nächstgelegene Schule gilt auch die Schule, die auf Anordnung der Schulbehörde besucht wird.  
Anträge auf Erstattung **sind** beim **Träger** der **Schülerbeförderung** **spätestens bis zum 30. September** eines jeden Jahres für das jeweils zurückliegende Schuljahr **einzureichen**.

Antrag auf Schülerbeförderung SchulG LSA

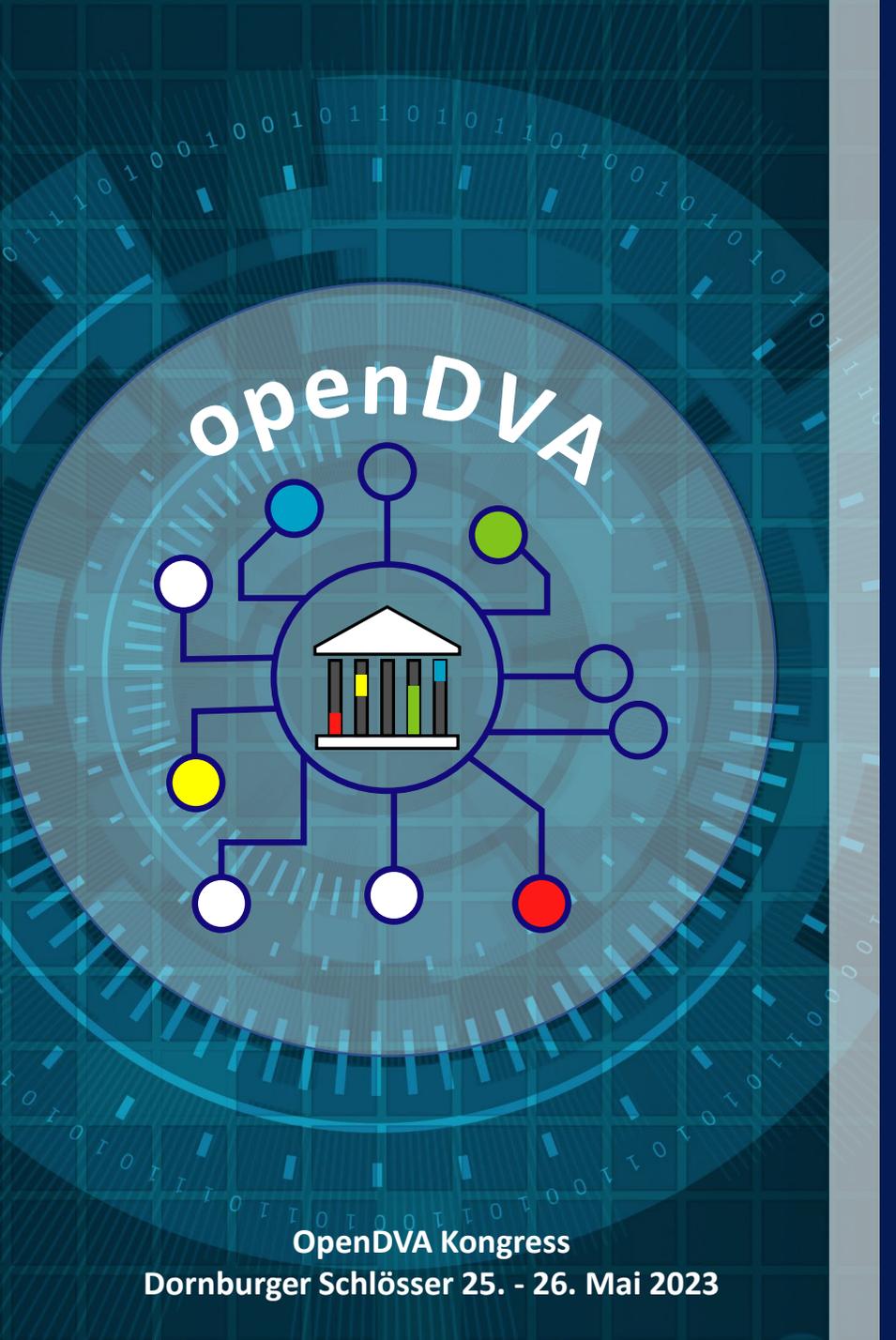
	Hauptakteur	Wer? (ausschließlich Verwaltungssicht)	} Rollen (Definiert FIM-Bausteine „Prozesselemente“)
	Initiator	Für wen?	
	Ergebnisempfänger	Für wen?	
	Mitwirkender	Mit wem?	
	Aktion	Was?	
	Signalwort	Welche Verbindlichkeit? (muss, soll, kann)	
	Ressource	Womit? z. B. Antrag (Name des Dokumentsteckbriefs)	
	Bedingung	Welche Voraussetzung? z. B. Zeitlicher Auslöser?	

# Ziel

Unterstützung der manuellen Normenanalyse durch automatisch generierte Vorschläge **und** stetige Verbesserung des Systems durch menschliche AnnotatorInnen



Human in the Loop



# Korpuserstellung

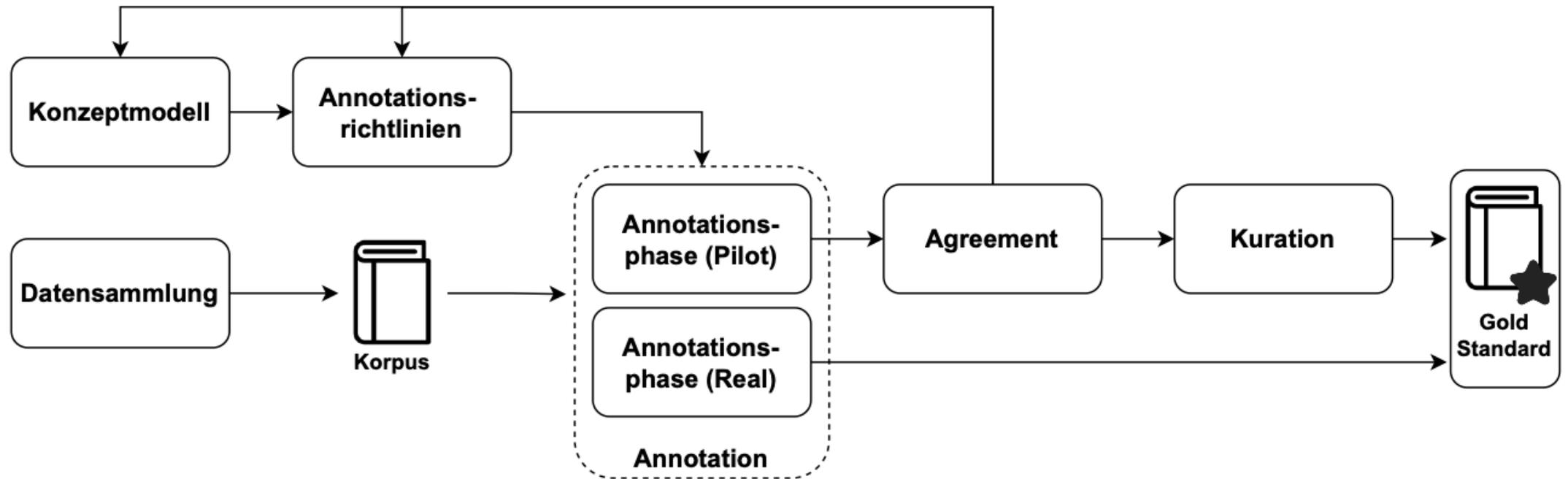
OpenDVA Kongress  
Dornburger Schlösser 25. - 26. Mai 2023



# Korpuserstellung

## Grund

- Keine vorhandenen Trainingsdaten
- Systematisierung der bisherigen Normenanalyserichtlinien

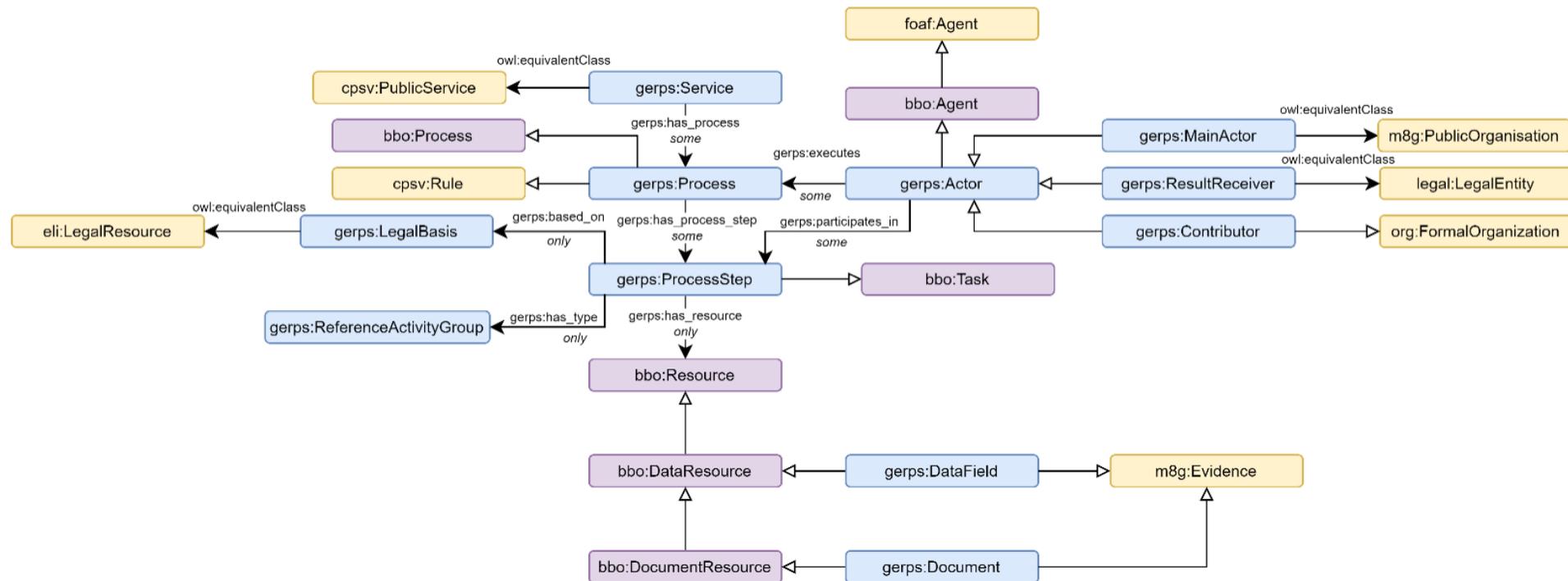
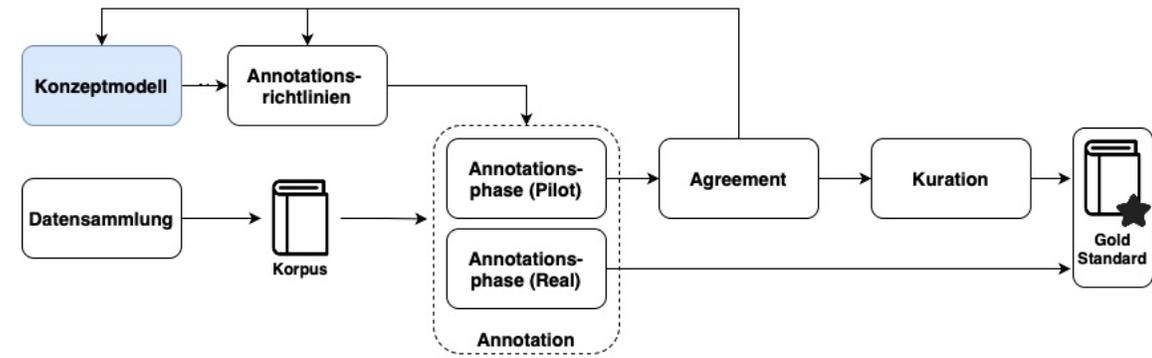




# Abgeschlossene Schritte

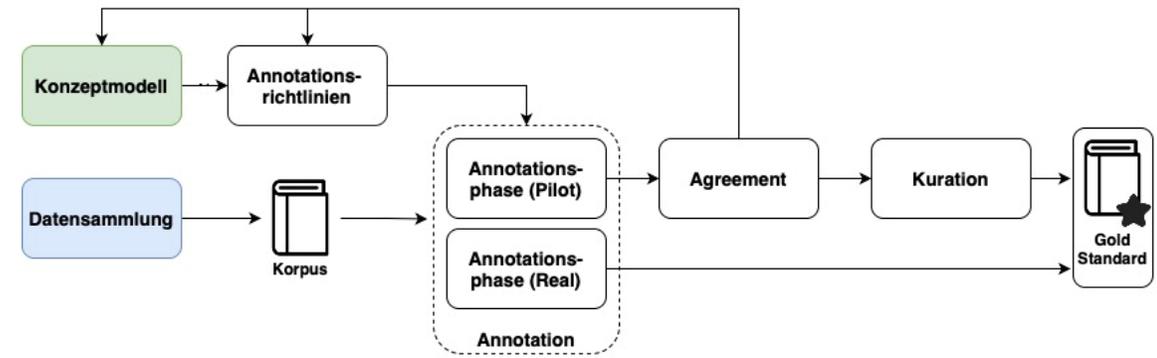
## Konzeptmodell

- Erstellung einer Ontologie (GerPS-onto)





# Abgeschlossene Schritte



## Konzeptmodell

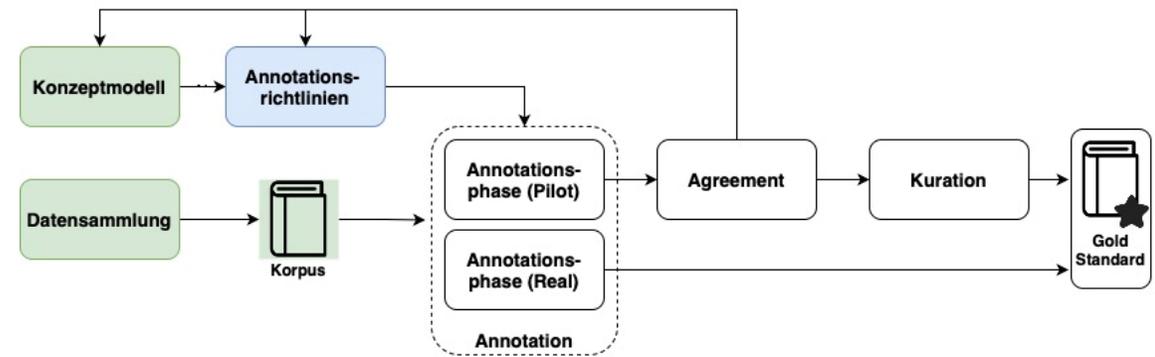
- Erstellung einer Ontologie (GerPS-onto)

## Datensammlung

- Benötigte Daten: Rechtsgrundlagen, aus denen sich Verwaltungsleistungen ableiten lassen
- Zu vorhandenen Leistungen im FIM-Portal, wurden Verweise zu Gesetzestexten extrahiert und automatisiert von [gesetze-im-internet.de](http://gesetze-im-internet.de) heruntergeladen.
- Für ein ausbalanciertes Korpus: Gleiche Anzahl an Texten für jede Leistungskategorie



# Abgeschlossene Schritte



## Konzeptmodell

- Erstellung einer Ontologie (GerPS-onto)

## Datensammlung

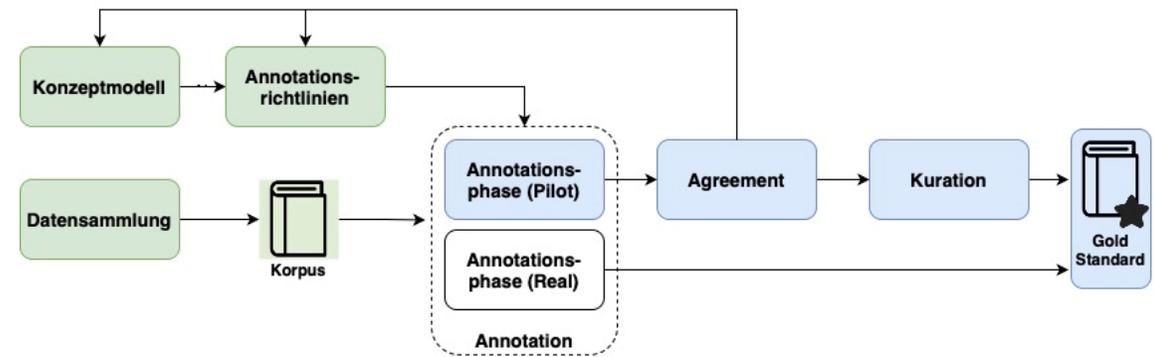
- Benötigte Daten: Rechtsgrundlagen, aus denen sich Verwaltungsleistungen ableiten lassen
- Zu vorhandenen Leistungen im FIM-Portal, wurden Verweise zu Gesetzestexten extrahiert und automatisiert von [gesetze-im-internet.de](http://gesetze-im-internet.de) heruntergeladen.
- Für ein ausbalanciertes Korpus: Gleiche Anzahl an Texten für jede Leistungskategorie

## Annotationsrichtlinien

- Erweiterung und Präzisierung der FIM-Kategorien



# Abgeschlossene Schritte



## Konzeptmodell

- Erstellung einer Ontologie (GerPS-onto)

## Datensammlung

- Benötigte Daten: Rechtsgrundlagen, aus denen sich Verwaltungsleistungen ableiten lassen
- Zu vorhandenen Leistungen im FIM-Portal, wurden Verweise zu Gesetzestexten extrahiert und automatisiert von [gesetze-im-internet.de](http://gesetze-im-internet.de) heruntergeladen.
- Für ein ausbalanciertes Korpus: Gleiche Anzahl an Texten für jede Leistungskategorie

## Annotationsrichtlinien

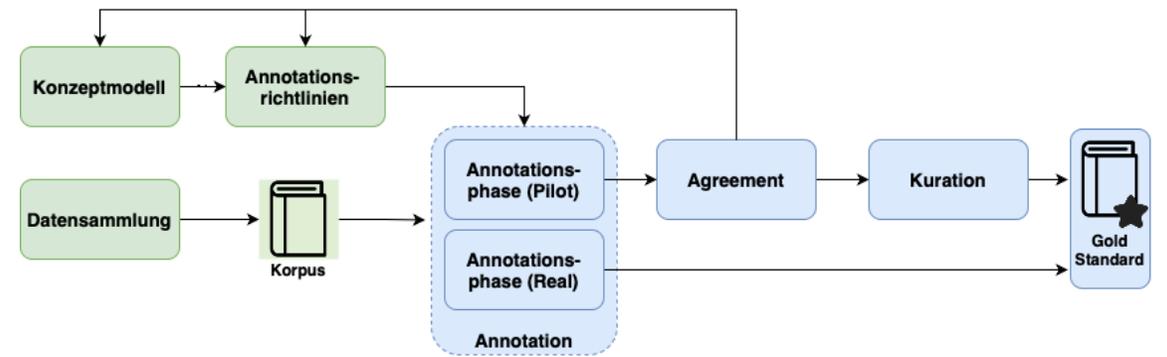
- Erweiterung und Präzisierung der FIM-Kategorien

## Erste Pilotphase

- 3 Annotatorinnen, 10 Gesetzestexte
- Iterative Anpassung der Annotationsrichtlinien



# Aktuelle Arbeitsphase

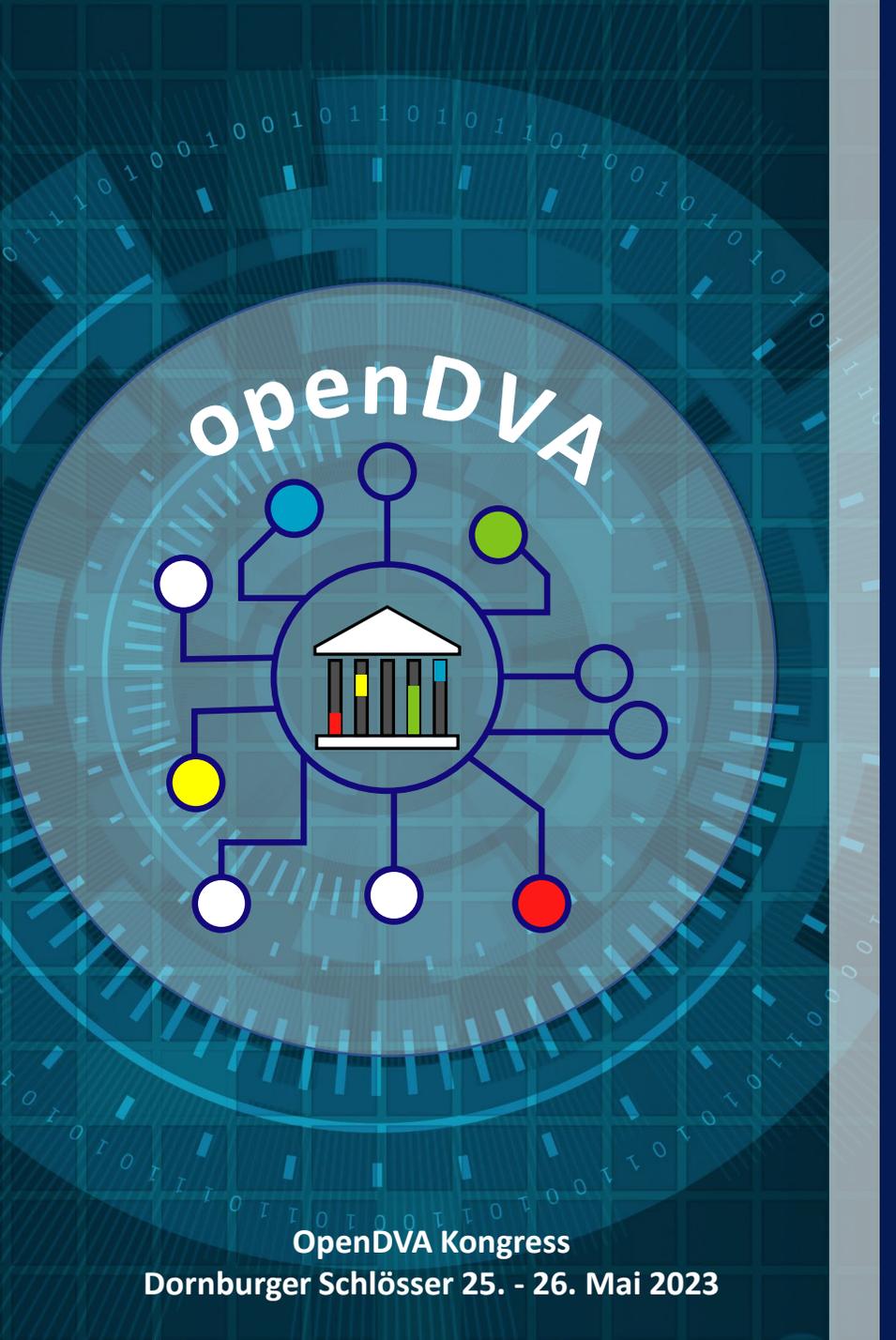


## Rahmen

- 3 Annotatoren aus der Rechts- und Verwaltungsdomäne
- 1010 zu annotierende Dokumente
- 3 Annotationsphasen

## Qualitätssicherung

- Schulung mit Hilfe der Annotationsrichtlinien
- Regelmäßige engmaschige Betreuung
- In den Anfangsphasen: Berechnung von Agreement-Messwerten zur Überprüfung der Übereinstimmung der Annotationen
- Bei großen und übereinstimmenden Bedenken seitens der Annotatoren erneute Anpassung der Richtlinien



# Vorstellung der Assistenzsysteme

OpenDVA Kongress  
Dornburger Schlösser 25. - 26. Mai 2023



# Ansätze für das Assistenzsystem

## Künstliche Intelligenz-Systeme

- Lernen durch die Analyse großer Datenmengen
- überwachtes Lernen: Algorithmen erhalten annotierte Daten, von denen sie lernen können
- Was gelernt werden muss: Wortgruppen Kategorien zuordnen
- Einige getestete Algorithmen: Bidirectional Long Short-Term Memory with a Conditional Random Field layer (BiLSTM-CRF) und XLM-RoBERTa (multilingual)

Die **Dokument** Erlaubnis ist zu **Aktion** befristen und hat im Falle der Nummer 1 Bestimmungen über **Datenfeld** Art , **Datenfeld** Umfang und **Datenfeld** Zeitpunkt des Eingriffs und die durchführende Person zu **Aktion** enthalten .

### Auszug XLM-R Modellvorhersagen



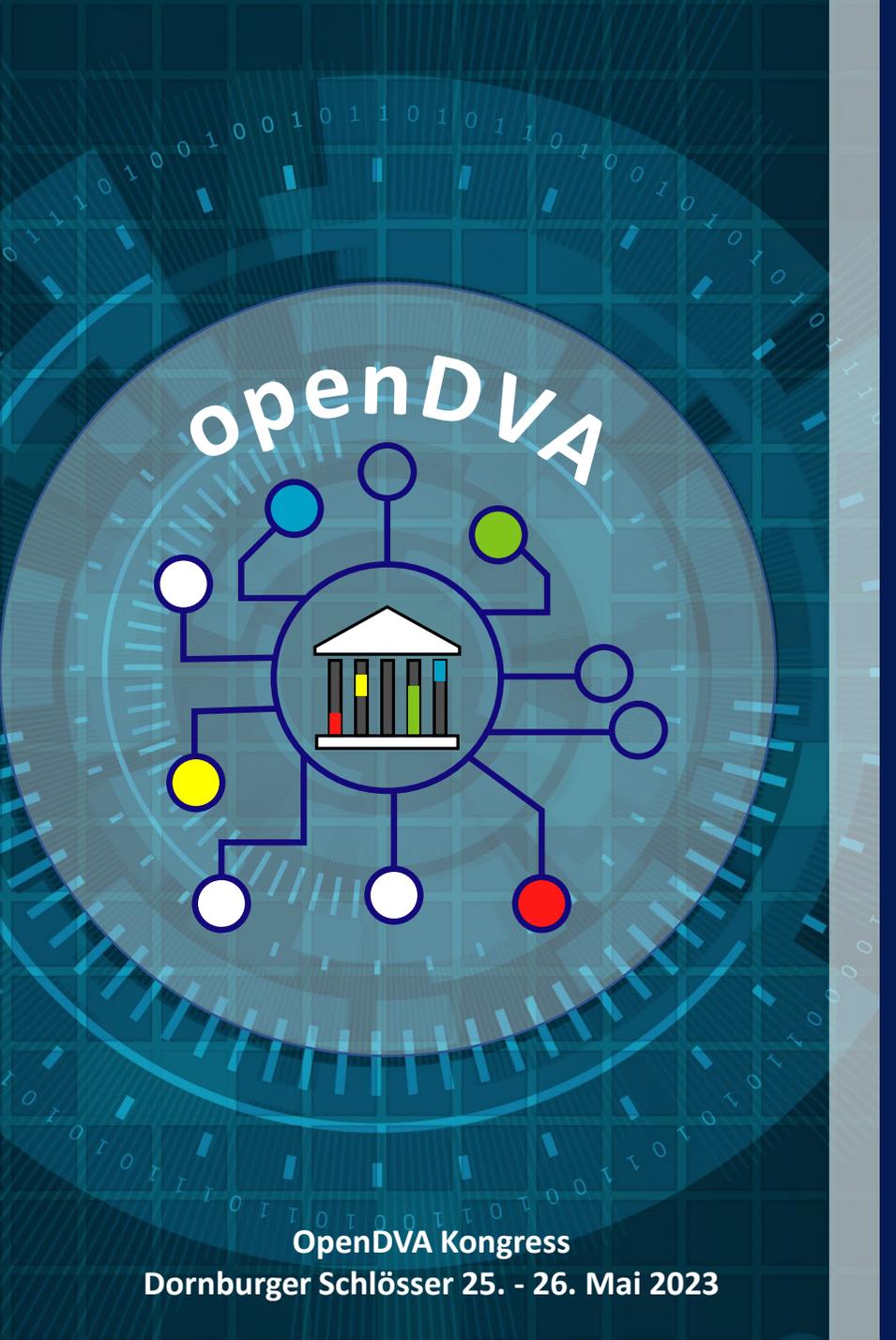
# Ansätze für das Assistenzsystem

## Künstliche Intelligenz-Systeme

- Lernen durch die Analyse großer Datenmengen
- überwachtes Lernen: Algorithmen erhalten annotierte Daten, von denen sie lernen können
- Was gelernt werden muss: Wortgruppen Kategorien zuordnen
- Einige getestete Algorithmen: Bidirectional Long Short-Term Memory with a Conditional Random Field layer (BiLSTM-CRF) und XLM-RoBERTa (multilingual)

## Regelbasierte Systeme

- Ableitung von Regeln vom existierenden Golddatenkorpus
  - Bestimmte Wortarten oder Wortbildungsmuster (Aktion)
  - Bestimmte Satzbildungsmuster (Bedingung)
- Erstellen von Gazetteers (Mitwirkender)
- Weiterverwendung von bestehenden Annotationslösungen wie LexNLP (Frist)



# Workshop: Intelligente und interaktive Normenanalyse

OpenDVA Kongress  
Dornburger Schlösser 25. - 26. Mai 2023





# Human-in-the-Loop-Tool in der Entstehung

History

- H** Akzeptiert Bundesamt
- S** Hinzugefügt kann

Annotationsvorschläge

**§ 9 BSI-G - Einzelnorm**

Gesetz über das Bundesamt für Sicherheit in der Informationstechnik (BSI-Gesetz - BSI-G) § 9 Zertifizierung

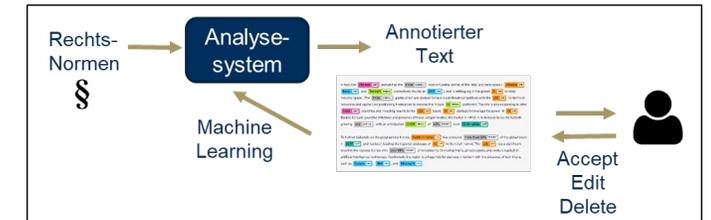
(1) Das Bundesamt ist nationale Zertifizierungsstelle der Bundesverwaltung für IT-Sicherheit .

(2) Für bestimmte Produkte oder Leistungen **Signalwort kann** beim **Hauptakteur Bundesamt** eine **Dokument Sicherheits-** und **Dokument Personenzertifizierung** oder eine **Aktion Zertifizierung** als **Aktion beantragt** werden. **Mitwirkender IT-Sicherheitsdienstleister**

Auswahl

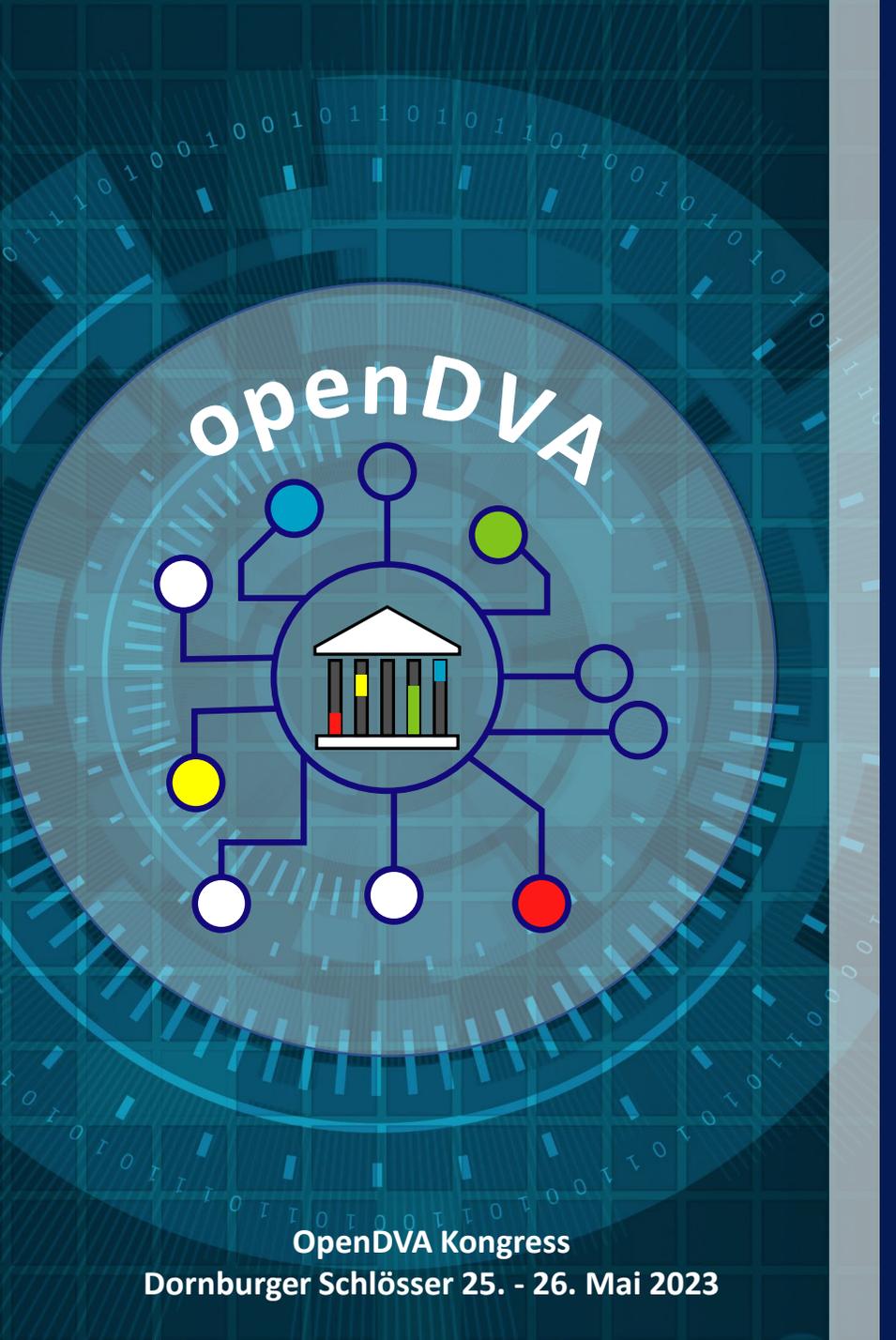
- Dokument
- Hauptakteur
- Ergebnisempfänger
- Aktion
- Mitwirkender
- Handlungsgrundlage
- Bedingung
- Signalwort

X ✓



**Workshop: Intelligente und interaktive Normenanalyse: 25.05. 15:30 - 17:00 !!**





# Ausblick

OpenDVA Kongress  
Dornburger Schlösser 25. - 26. Mai 2023



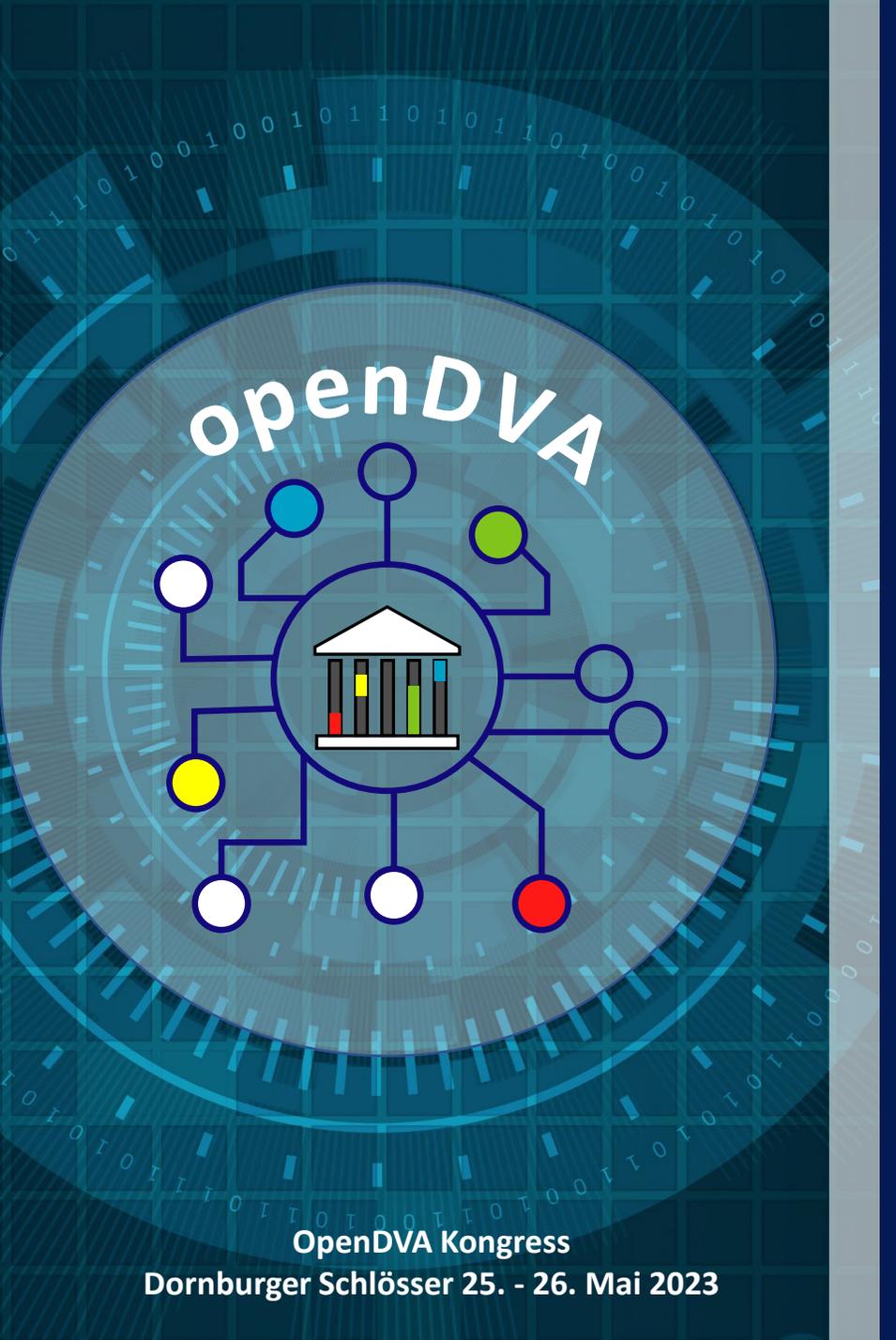
# Herausforderungen

- Geringe Menge annotierter Daten
- Uneinheitliche Kategorien
  - Bedingung: Nebensätze, Aktion: Verben und Nomen, Signalwörter: Modalverben
  - Nur Hauptakteur, Mitwirkender und manchmal Ereignisempfänger sind Eigennamen
- Unterscheidung zwischen verschiedenen Akteuren
- Bestimmung administrativer Aktionen.
- Unterscheidung zwischen Relevantem und Irrelevantem für die Leistungen



# Nächste Schritte

- Erweiterung des Trainingskorpus
- Iteratives Trainieren und Testen der Künstliche Intelligenz-Systeme
- Weiterentwicklung der regelbasierten Ansätze
- Zusammenfassung der Annotationsergebnisse -> wie können wir vom annotierten Text Prozessschritte ableiten?
- Fortführung der Toolentwicklung und -designs
- Evaluation des Tools durch Endnutzer



# Code und Links

OpenDVA Kongress  
Dornburger Schlösser 25. - 26. Mai 2023



# Materialsammlung

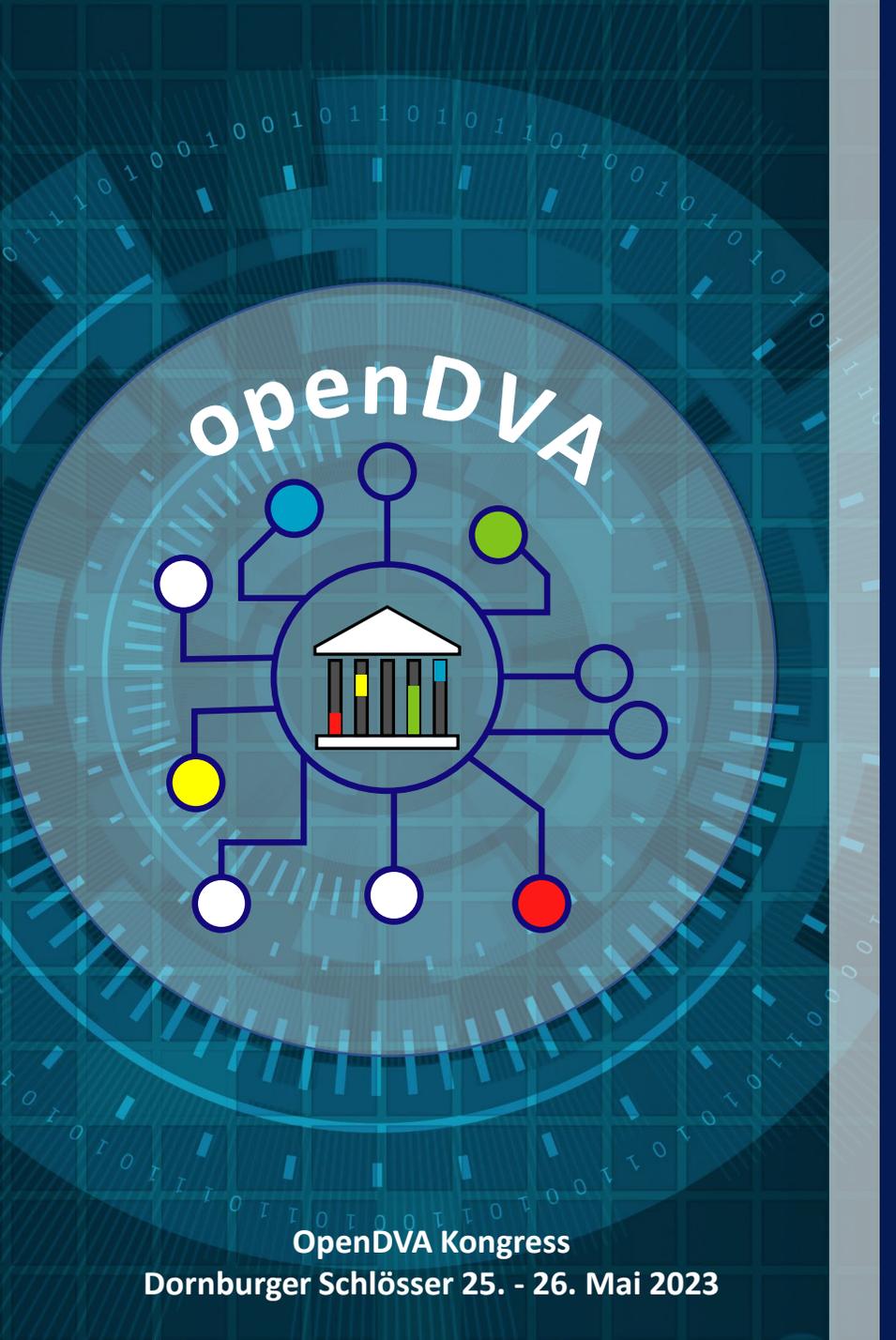
## Ontology

- Code: <https://github.com/fusion-jena/GerPS-onto>
- Doku: <https://fusion-jena.github.io/gerps-ontology/index.html>

## Datensammlung

- Code: <https://doi.org/10.5281/zenodo.7875288>
- Daten: <https://doi.org/10.5281/zenodo.7900297>





# Diskussion

OpenDVA Kongress  
Dornburger Schlösser 25. - 26. Mai 2023





**Vielen Dank  
für Ihre Aufmerksamkeit!**

Kontakt:

[leila.feddoul@uni-jena.de](mailto:leila.feddoul@uni-jena.de)  
[clara.lachenmaier@uni-Bielefeld.de](mailto:clara.lachenmaier@uni-Bielefeld.de)

Gefördert durch:



Bundesministerium  
des Innern  
und für Heimat

aufgrund eines Beschlusses  
des Deutschen Bundestages

## Kollaboration:

