

# Wie robust ist die vollautomatische Anonymisierung von Gerichtsentscheidungen?

**Bao Minh Doan Dang Julian Werner**

Lehrstuhl für Korpus- und Computerlinguistik  
Friedrich-Alexander-Universität Erlangen-Nürnberg



Funded by  
the European Union  
NextGenerationEU

With funding from the:



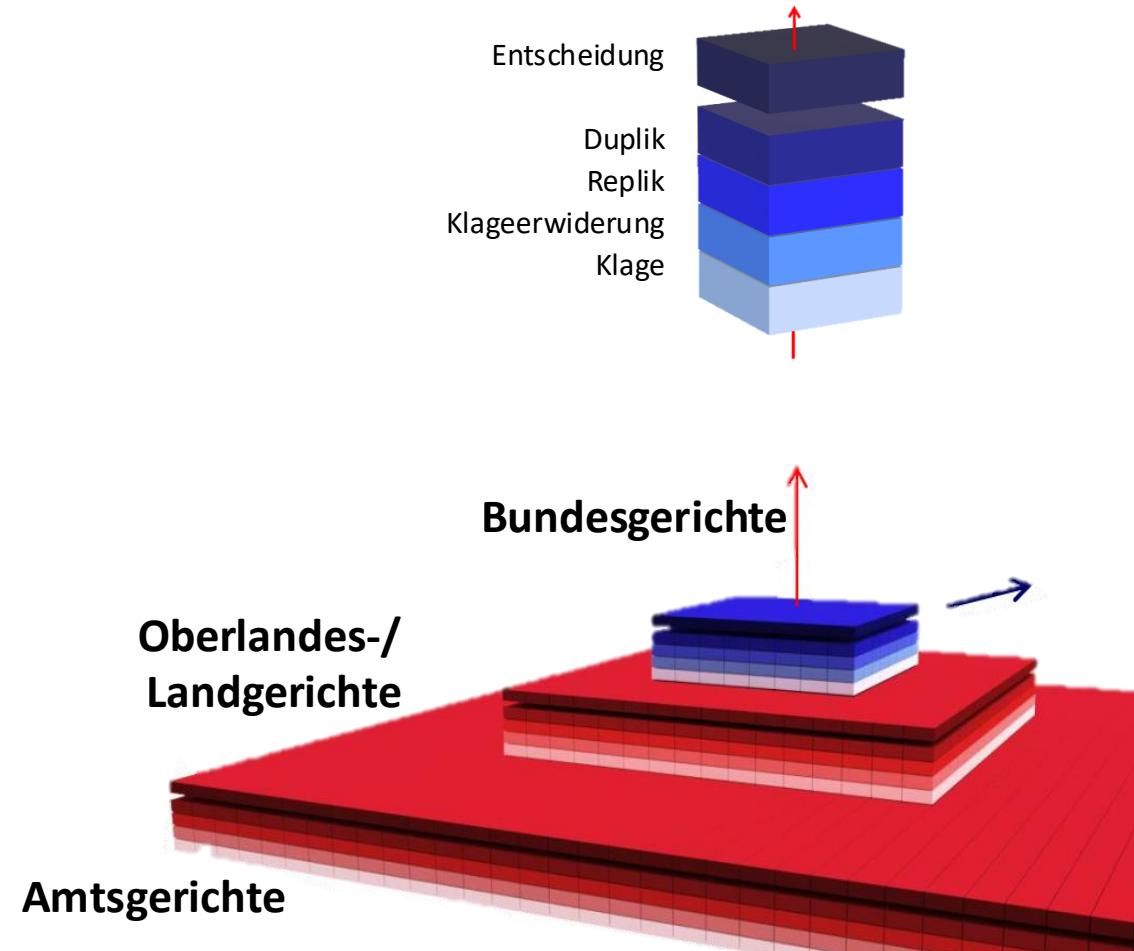
Federal Ministry  
of Research, Technology  
and Space

# Agenda

- 1. Motivation**
- 2. Goldstandard**
- 3. Automatische Anonymisierung**
- 4. Fazit & Ausblick**

# 1. Motivation

- Gesetzliche Verpflichtung der Gerichte zur Veröffentlichung (aller) Urteile
  - Gesamtvolumen:  
ca. 1,5 Millionen Urteile / Jahr
  - Grundrecht auf informationelle Selbstbestimmung → Urteile müssen für die Veröffentlichung anonymisiert werden
- Aktuelle Situation: < 3% der Urteile werden veröffentlicht, meist von Obergerichten und Oberlandesgerichten, weil die manuelle Anonymisierung zu teuer ist
- Erstinstanzliche Entscheidungen sind unterrepräsentiert, d.h. diejenigen, die den Sachverhalt eines Falles erörtern, wären die nützlichsten Daten für Legal Tech



## Relevante Rechtsvorschriften:

- Recht auf informationelle Selbstbestimmung (Art. 2 Abs. 1 GG)
- DSGVO (schützt nur natürliche Personen, jur. Personen werden nicht erfasst)
- § 30 Abgabenordnung (Steuergeheimnis)
- §§ 203 ff. StGB
  - → eine erfolgreiche Anonymisierung muss sicherstellen, dass eine Re-Identifizierung einen unangemessen hohen Aufwand erfordern würde (Zeit, Kosten, Technologieeinsatz etc.)
  - OLG Karlsruhe vom 22.12.2020, 6 VA 24/20
  - VGH Baden-Württemberg vom 23.7.2010, 1 S 501/10

## Direkte Identifikatoren (PII = personally identifying information):

- Name (natürliche und juristische Personen)
- Adresse, Telefonnummer, Kfz-Kennzeichen, etc.
- Geburtsdatum usw.

## Mittelbar identifizierende Informationen:

- Angaben zu Beruf bzw. Ausbildung
- Akademischer Titel
- Gesundheitsdaten
- deskriptive Angaben
- besondere Merkmale (z.B. einziges rotes Haus in einer kleinen Ortschaft)
- → **Deanonymisierung durch Verknüpfung**

- Anonymisierung als Vorgang: Verknüpfbarkeit von Merkmalsausprägungen mit einem Individuum verunsichert durch Vergrößerung der Menge an möglichen Merkmalsträgern dieser Ausprägungen
- Für die (indirekte) Identifizierbarkeit sind alle objektiven Faktoren wie die Kosten der Re-Identifizierung und der dafür erforderliche Zeitaufwand und Arbeitseinsatz zu berücksichtigen
- Maßgeblich: zum Zeitpunkt der Anonymisierung verfügbare Technologie und die technologischen Entwicklungen (Erwg. 15, 26 DSGVO).

- Beispiel Wohnort: nicht nur über die Adresse darstellbar, sondern über deskriptive Merkmale:
- Das(einzig)e rote Haus in einem bestimmten kleinen Dorf
- → Hier lauert eine häufig unterschätzte Gefahr in Urteilen aufgrund von kostengünstigen Verknüpfungsmöglichkeiten eine De-Anonymisierung herbeizuführen
- Anonymität ist dann gewährleistet, wenn Zeit-, Kosten- und Technik-Aufwand zu groß, als dass mit Re-Identifizierung zu rechnen ist (OLG Karlsruhe vom 22.12.2020, 6 VA 24/20; VGH Baden-Württemberg vom 23.7.2010, 1 S 501/10)
- Verfahren zur De-Anonymisierung und daraus resultierende Gefahren einer Re-Identifizierung sind noch unzureichend erforscht. Es sind empirische Daten zum Arbeits- und Zeiteinsatz sowie zum verfügbaren Zusatzwissen notwendig.

## Direkte Identifikatoren (PII = personally identifying information):

- Name (natürliche und juristische Personen)
- Adresse, Telefonnummer, Kfz-Kennzeichen, etc.
- Geburtsdatum usw.

## Mittelbar identifizierende Informationen (Pseudoidentifikatoren):

- Angaben zu Beruf bzw. Ausbildung
- Akademischer Titel
- Gesundheitsdaten
- deskriptive Angaben (örtliche Verhältnisse, Betriebsinformationen)
- einzigartige Merkmale (z.B. einziges rotes Haus in einer kleinen Ortschaft)

- **Formales**
  - Aktenzeichen
  - Gericht
- **Natürliche Person**
  - Name
  - Juristische Funktionsträger
  - Sonstiges identifizierendes Merkmal
- **Juristische Person**
  - Name
  - Sonstiges identifizierendes Merkmal
- **Adresse**
  - Ortsangabe
  - Sonstiges identifizierendes Merkmal
- **Datum**
  - Weltwissen
  - Sachverhalt
  - Prozessgeschichte
- **Fahrzeug**
  - Kennzeichen
  - Fahrgestellnummer
  - Sonstiges identifizierendes Merkmal
- **Immateriagut**
  - Name
  - Sonstiges identifizierendes Merkmal
- **Sonstiges**
  - Kontaktdaten
  - Geschäftszeichen

Weitere Optionen bei allen Tags:

- Risikoneuau (hoch/mittel/niedrig)
- Informationserhaltung notwendig
- unsicher/diskussionswürdig

## 1. Manuelle Anonymisierung

- unterstützt durch Texteditor-Plugins („office tech“)

## 2. Halbautomatische Anonymisierung

- manuelle Überprüfung automatisch erzeugter Vorschläge → umgeht KI-Verordnung
- selbstlernende KI verbessert die automatischen Vorschläge durch menschliches Feedback

## 3. Vollautomatische Anonymisierung

- nur dieser Ansatz skaliert auf 1,5 Mio. Urteile/Jahr
- bisherige Lösungsansätze auf Basis von Named Entity Recognition (NER) lassen die meisten Pseudoidentifikatoren außer Acht

- **Existierende Tools:**

- [A-Tool](#) (BALO.AI)
  - Plugin für MS Word, im Einsatz bei Schweizer Gerichten
- EU-gefördertes Projekt [MAPA](#)
  - [Vollautomatische Anonymisierung](#) mit NER-Ansatz
- [OpenRedact](#) (BMBF Prototype Fund)
  - Halbautomatisches Open-Source Tool mit angepasster NER
- [Text Anonymization Benchmark \(TAB\)](#)
  - Goldstandard: EGMR-Urteile mit semi-automatischer Annotation
- [HILANO](#) (BMBF kmu-Innovativ)
  - Selbstlernende KI mit *human-in-the-loop*-Ansatz
- [JANO](#) (IBM für Hessen & Baden-Württemberg)
  - halbautomatisch: Empfehlungen für manuelle Anonymisierung

## LeAK

- gefördert 2020–2023 durch BayStMJ
- Team: Nathan Dykes, Philipp Heinrich, Michael Keuchen, Thomas Proisl  
+ studentische Annotator:innen

## AnGer

- gefördert 2023–2025 durch BMBF/NextGenerationEU
- Team: Bao Minh Doan Dang, Philipp Heinrich, Michael Keuchen, Mahdi Mantash, Melanie Rosa, Naveed Unjum, Julian Werner, Leonardo Zilio  
+ zahlreiche studentische Annotator:innen

### Ziele:

- Machbarkeitsstudie zur vollautomatischen Anonymisierung von Gerichtsentscheidungen mit Schwerpunkt auf Amtgerichten (AG) und zwei Rechtsgebieten (Mietrecht, Verkehr)
- Bestimmung des Deanonymisierungsrisikos für verschiedene Maskierungstechniken

### Ziele:

- Weiterentwicklung zu einem anwendungsreifen Prototyp durch Finetuning vortrainierter LLM
- für alle Instanzen und über viele Rechtsgebiete hinweg
- Domänenanpassung, Data Augmentation, etc.
- weitere Experimente zum Deanonymisierungsrisiko

- Forschungsziel von AnGer: Weiterentwicklung des Anonymisierungstoolkits für neue Domänen
- Domänen im Sinne des Projekts: Rechtsgebiete und Instanzen
- LeAK: Amtsgericht → Verkehrsrecht und Mietrecht
- AnGer: Oberlandesgericht → 11 neue Rechtsgebiete
- Robustheit in NLP: Ein robustes Modell erzielt bei unbekannten Domänen weiterhin konsistent verlässliche Ergebnisse (**Generalisierung**)
- Robustheit in KI-VO?

### Genauigkeit, Robustheit und Cybersicherheit

- (1) Hochrisiko-KI-Systeme werden so konzipiert und entwickelt, dass sie ein angemessenes Maß an Genauigkeit, **Robustheit** und Cybersicherheit erreichen und in dieser Hinsicht während ihres gesamten Lebenszyklus beständig funktionieren.
- (2) Um die technischen Aspekte der Art und Weise der Messung des angemessenen Maßes an Genauigkeit und **Robustheit** gemäß Absatz 1 und anderer einschlägiger Leistungsmetriken anzugehen, fördert die Kommission in Zusammenarbeit mit einschlägigen Interessenträgern und Organisationen wie Metrologie- und Benchmarking-Behörden gegebenenfalls die Entwicklung von Benchmarks und Messmethoden.
- (3) Die Maße an Genauigkeit und die relevanten Genauigkeitsmetriken von Hochrisiko-KI-Systemen werden in den ihnen beigefügten Betriebsanleitungen angegeben.
- (4) Hochrisiko-KI-Systeme müssen so widerstandsfähig wie möglich gegenüber Fehlern, Störungen oder Unstimmigkeiten sein, die innerhalb des Systems oder der Umgebung, in der das System betrieben wird, insbesondere wegen seiner Interaktion mit natürlichen Personen oder anderen Systemen, auftreten können. <sup>2</sup>In diesem Zusammenhang sind technische und organisatorische Maßnahmen zu ergreifen.

**Die Robustheit von Hochrisiko-KI-Systemen kann durch technische Redundanz erreicht werden, was auch Sicherungs- oder Störungssicherheitspläne umfassen kann.**

[...]

### Resilienz von Hochrisiko-KI-Systemen

<sup>1</sup>Die technische Robustheit ist eine wesentliche Voraussetzung für Hochrisiko-KI-Systeme. <sup>2</sup>Sie sollten widerstandsfähig in Bezug auf schädliches oder anderweitig unerwünschtes Verhalten sein, das sich aus Einschränkungen innerhalb der Systeme oder der Umgebung, in der die Systeme betrieben werden, ergeben kann (z. B. Fehler, Störungen, Unstimmigkeiten, unerwartete Situationen). <sup>3</sup>Daher sollten technische und organisatorische Maßnahmen ergriffen werden, um die Robustheit von Hochrisiko-KI-Systemen sicherzustellen, indem beispielsweise geeignete technische Lösungen konzipiert und entwickelt werden, um schädliches oder anderweitig unerwünschtes Verhalten zu verhindern oder zu minimieren. [...]

## Zweckbestimmung und Datenauswahl

---

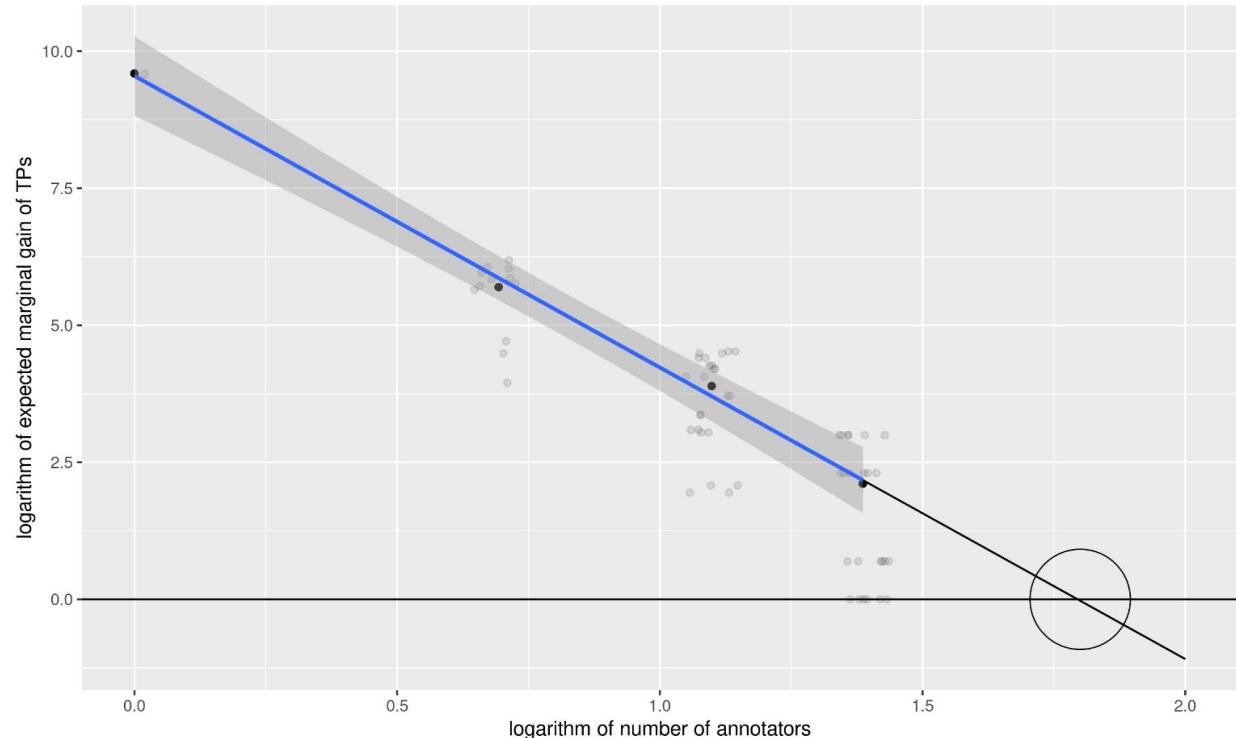
- Art. 3 Nr. 12 KI-VO definiert Zweckbestimmung
  - Anbieter muss Einsatzgebiet und Umstände und Bedingungen für die Verwendung angeben
- Flankierend: Art. 10 KI-VO bestimmt, dass bei Trainings- und Evaluationsdatenauswahl entsprechend der Zweckbestimmung die korrekten Daten ausgewählt werden müssen
- Domänenanpassung also besser abgebildet über Zweckbestimmungserfordernis in KI-VO

## 2. Goldstandard

- **EU-KI-Verordnung:** automatische Anonymisierung muss nachweislich zweckmäßig, genau, robust sein.
- Höchste Anforderungen an die Zuverlässigkeit von automatischen Systemen:
- 95% Genauigkeit in NLP und KI hervorragend, skaliert aber nicht auf 1,5 Mio Urteile:  
5% Fehler → 75.000 DSGVO-Fälle / Jahr
- Genauigkeit wird in NLP oft pro Token gemessen: System wird dann auch für nicht zu maskierende Wörter „belohnt“, selbst wenn es viele kritische Textstellen übersieht
- Konsequenz: > 99% Recall für PII erforderlich
- weniger strenge Anforderungen an Pseudo-identifikatoren (je nach ihrer Selektivität)

## Goldstandard wird benötigt

- für Training (z.B. mit Deep Learning)
- für Evaluation → Nachweis der Zuverlässigkeit
- Goldstandard wird durch manuelle Annotation erstellt



## Extrem hohe Qualität des Goldstandards

- kann System nicht auf > 99% Recall testen, wenn der Goldstandard selbst nur zu 97% korrekt ist (in NLP üblich, bei TAB wahrscheinlich noch niedriger)
- Goldstandard muss praktisch fehlerfrei sein!
- empirisches Modell mit mathematischer Extrapolation → jeder Text muss von 5–6 Annotatoren bearbeitet werden (Heinrich et al. 2021)

**Goldstandard** bedeutet, dass jeder Urteilstext mindestens

- **5-fach annotiert** und
- **zusätzlich adjudiziert** wurde

→ damit fast sicher alle kritischen Textstellen gefunden

Alle annotierten Textstellen wurden durch realistische **Pseudonyme** ersetzt, um den Goldstandard auch für Experimente mit Deep-Learning-Verfahren auf HPC-Rechnern außerhalb des geschützten Raumes nutzen zu können

(NB: keine Pseudonymisierung im Sinne der DSGVO, da keine Zuordnung zu ursprünglichen Angaben → Surrogate)

## Goldstandards in diesem Vortrag:

- Amtsgericht (**AG**): Miet- & Verkehrsrecht  
570 Entscheidungen, ca. 1 Mio. Token
- Oberlandesgericht (**OLG**): 11 Rechtsgebiete  
1229 Entscheidungen, ca. 4.1 Mio. Token

Rechtsgebiet	Urteile	Tokens
Allg. Zivilsachen	236	848K
Bankensachen	177	375K
Bausachen	77	238K
Beschwerdeverfahren	46	101K
Familiensachen	34	84K
Handelssachen	174	789K
Immaterialgüter	108	559K
Kapitalschanlagen	129	468K
Kostensachen	22	39K
Schiedssachen	33	116K
Verkehrsunfallsachen	194	507K

### 3. Automatische Anonymisierung

# Ergebnisse: AG (Amtsgericht)

Training und Evaluation auf AG-Goldstandard (AnGer)

Rechtsgebiet	Precision	Recall	Recall (PII)
Mietrecht	96.20%	95.74%	98.90%
Verkehrsrecht	97.23%	97.46%	99.34%
<b>Gesamtergebnis</b>	<b>96.77%</b>	<b>96.73%</b>	<b>99.16%</b>

- Vortrainiertes Sprachmodell: deutsches **GottBert**-Modell (125 Mio. Parameter)
- Finetuning für die Erkennung von Textstellen (via BIO-Tagging) über 10 Epochen
- Die Lernrate wird mittels Grid Search (auf Development-Daten) optimiert
- **Multitask-Modell zur Erkennung von Textstelle, Informationskategorie und Risiko**
- 50% Trainingsdaten, 25% Development-Daten, 25% Testdaten

# Übertragung auf OLG-Urteile

Wie gut ist die Übertragbarkeit auf anderen Dokumenttyp bzw. Rechtsgebiet?

Rechtsgebiet	Precision	Recall	Recall (PII)
Allgemeine Zivilsachen	87.35%	88.40%	94.82%
Bankensachen	89.58%	94.47%	97.35%
Bausachen	88.10%	95.58%	96.53%
Beschwerdeverfahren	78.47%	93.55%	96.08%
Familiensachen	85.60%	91.23%	95.87%
Handelssachen	90.25%	93.65%	96.84%
Immaterialgüter	78.77%	76.38%	83.38%
Kapitalsachanlagen	87.55%	91.29%	96.55%
Kostensachen	82.80%	93.61%	95.0%
Schiedssachen	86.89%	91.25%	96.93%
Verkehrsunfallsachen	85.81%	92.42%	97.09%
Gold Standard	86.76%	90.52%	95.25%

**Distribution shift:** Keine zuverlässliche Übertragung auf OLG-Urteile trotz eines hohen in-domain Recalls > 99%

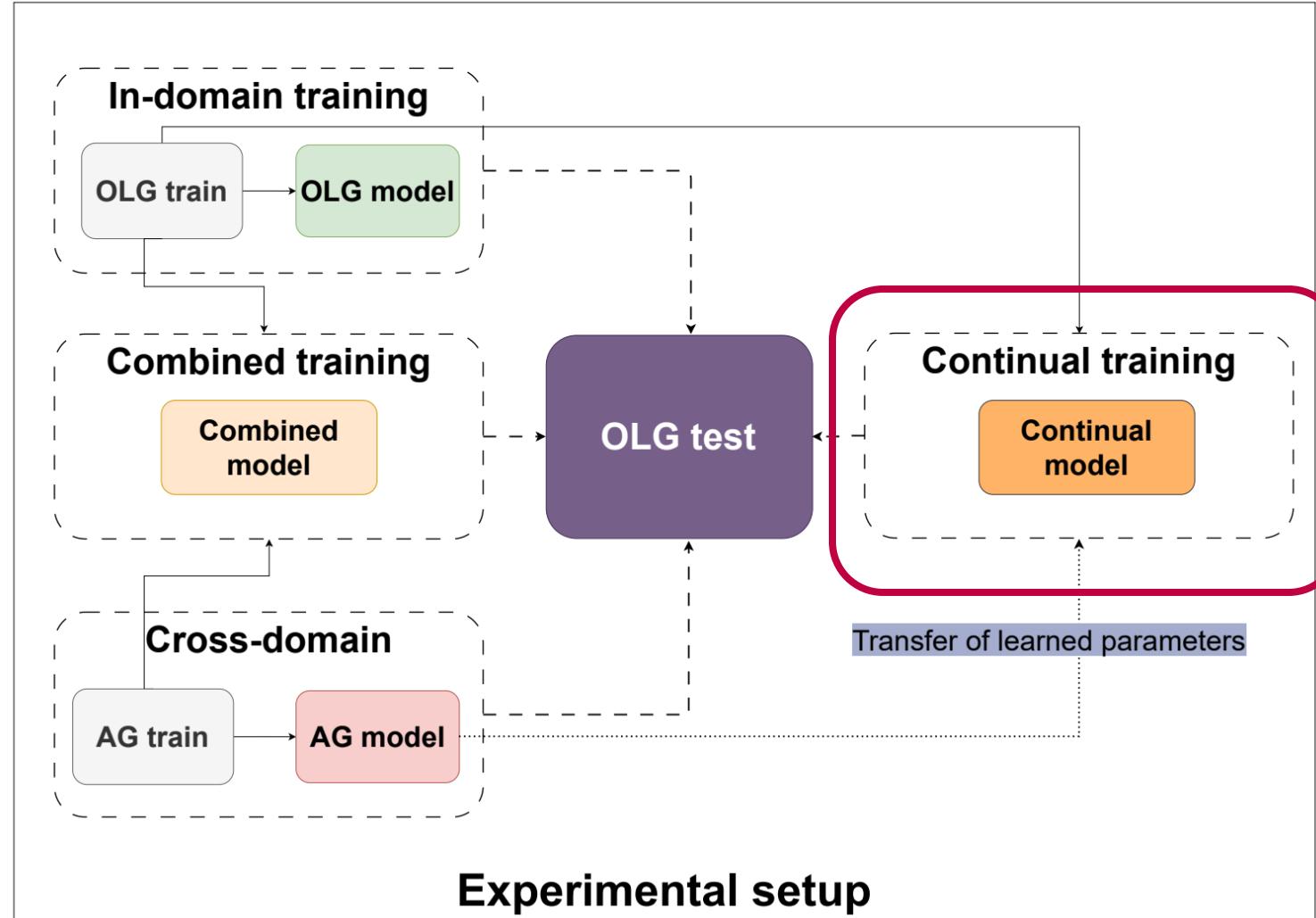
# Kann ein Data Augmentation Ansatz helfen?

Generierung von synthetischen Daten mit **MLM** (Masked Language Model)

Rechtsgebiet	Precision	Recall	MLM-Aug.	Recall (PII)	MLM-Aug. (PII)
Mietrecht	96.20%	95.74%	96.12%	98.90%	98.79%
Verkehrsrecht	97.23%	97.46%	97.16%	99.34%	99.48%
<b>Gesamtergebnis</b>	<b>96.77%</b>	<b>96.73%</b>	<b>96.70%</b>	<b>99.16%</b>	<b>99.21%</b>

- Grundlage: **MulDa** (Liu et al. 2021) und **MELM** (Zhou et al. 2022)
- Idee: Domänenanpassung eines vortrainierten Sprachmodells zur Erzeugung synthetischer Trainingsdaten
- Nur Wörter aus den zu anonymisierenden Textstellen werden zufällig maskiert
- Datengenerierung: Top-k Strategie (k=10) und eine Temperatur von 1

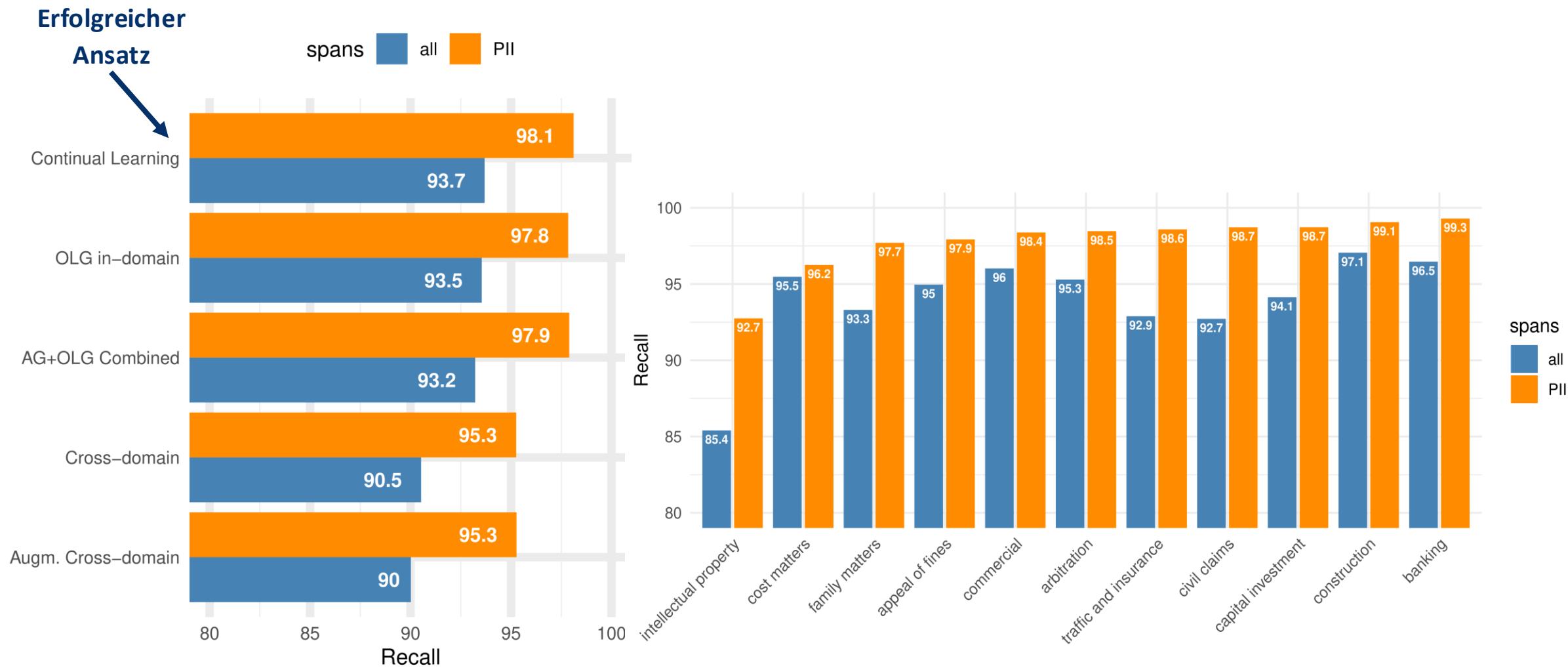
AG-Modell wird mit zusätzlichen Trainingsdaten aus OLG-Goldstandard optimiert



- Das AG-Modell wird auf OLG-Daten weiter angepasst
- Die gelernten Parameter bleiben erhalten
- Effiziente Anpassung an neue Daten und Domänen
- Domänenanpassungen wurden mit einer 5-fachen Kreuzvalidierung durchgeführt

# Experimentsergebnisse

Vergleich der verschiedenen Experimente mit OLG-Urteilen

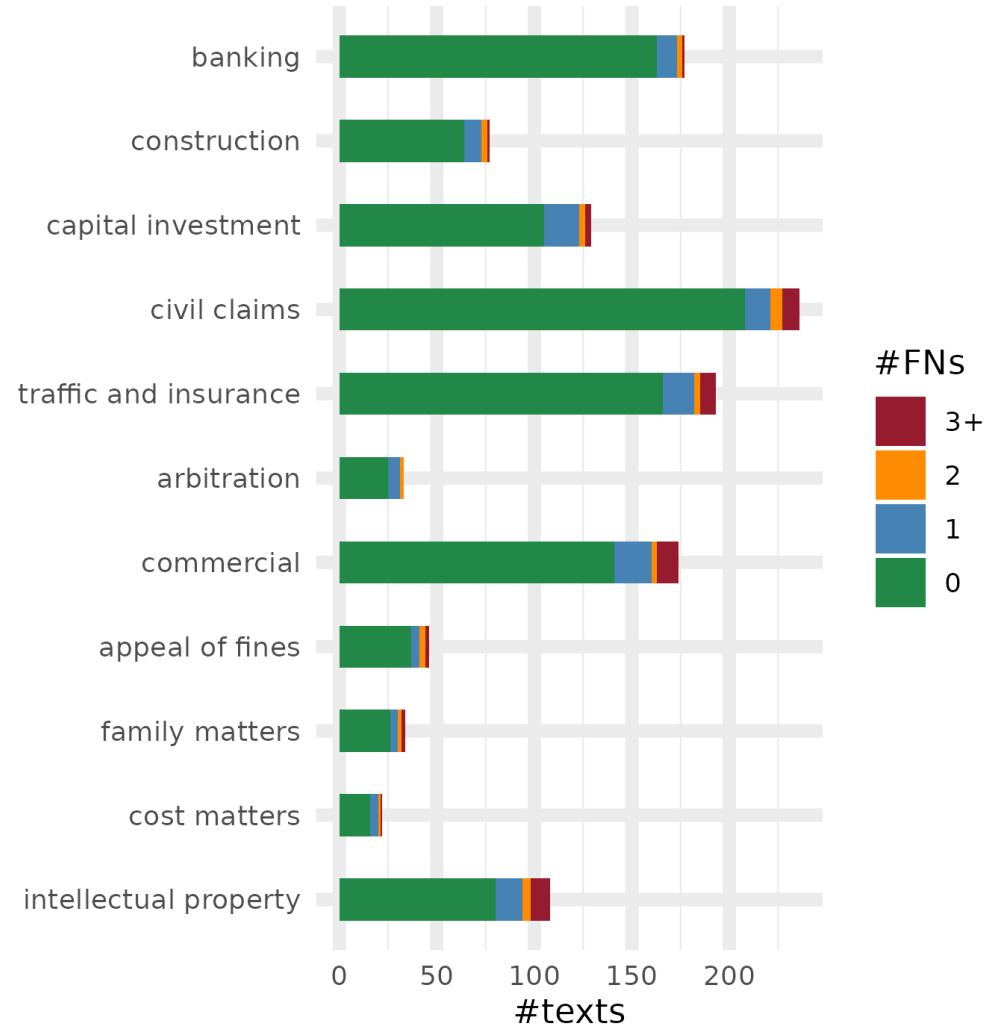


Rechtsgebiet	Precision	Recall	Continual	Recall (PII)	Continual (PII)
Allgemeine Zivilsachen	87.35%	88.40%	92.72%	94.82%	98.72%
Bankensachen	89.58%	94.47%	96.47%	97.35%	99.29%
Bausachen	88.10%	95.58%	97.07%	96.53%	99.05%
Beschwerdeverfahren	78.47%	93.55%	94.95%	96.08%	97.93%
Familiensachen	85.60%	91.23%	93.30%	95.87%	97.70%
Handelssachen	90.25%	93.65%	96.01%	96.84%	98.37%
Immaterialgüter	78.77%	76.38%	85.38%	83.38%	92.75%
Kapitalschanlagen	87.55%	91.29%	94.13%	96.55%	98.72%
Kostensachen	82.80%	93.61%	95.48%	95.0%	96.25%
Schiedssachen	86.89%	91.25%	95.29%	96.93%	98.46%
Verkehrsunfallsachen	85.81%	92.42%	92.89%	97.09%	98.58%
Gold Standard	86.76%	90.52%	93.68%	95.25%	98.11%

# Ergebnisse: OLG nach Informationskategorie

Korrekte Erkennung von Textstelle und Informationskategorie (vereinfachtes Schema)

Informationskategorie	Precision	Recall	# Textstellen
Personennamen	99.34%	99.15%	11448
Adresse	97.40%	98.21%	9996
Kfz-Kennzeichen	98.83%	97.38%	343
Firmennamen	97.86%	92.20%	8706
Pseudoidentifikatoren	95.67%	53.00%	8931
Datum	98.67%	83.84%	11384
Immateriagut	99.11%	76.77%	3268
Aktenzeichen (Gericht)	99.70%	92.97%	6507
PII (Person, Adresse, Kennz.)	98.52%	98.24%	21787



## Fehleranalyse für 117 Urteile mit exakt einem FN:

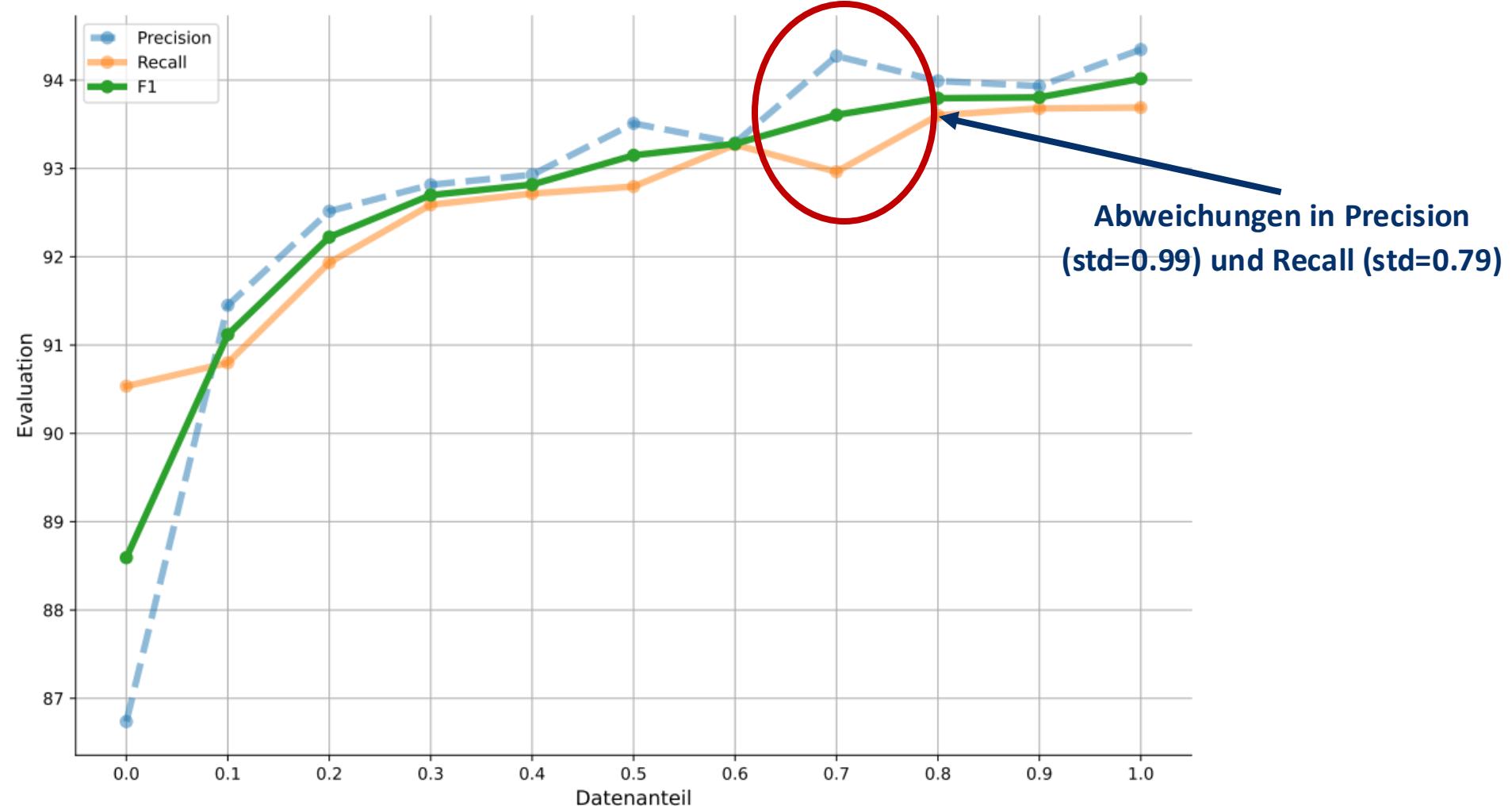
- 28 Textstellen wurden vom Modell aufgeteilt
- 15 Textstellen wurden vom Modell zusammengeführt
- 3 Überlappungen zwischen Goldstandard und Vorhersagen
- 5 Sonderzeichen wurden nicht korrekt erkannt
- 24 Fehler liegen im Goldstandard (Kein tatsächliches hohes Risiko)
- 23 Textstellen waren lediglich Pseudoidentifikatoren
- **19 tatsächliche FNs** (=16.2% von 117 Dokumenten)

## Beispiele:

- [[Esser, Dieter]], [[E. T. A. ] ]–Hoffmann-Straße 31d, 82496 Oberau].
- Denn [Frau [Lange]] behielt sich durch das Wohnrecht ...
- ... als Single veröffentlichte [Titel [Slow Life]] ...
- ... dass er heute zum [Baldauf] geht ...
- ... Voraussetzungen im Falle [Korte] unstreitig vorgelegen ...
- [Hartwig, F ] et al.: Mitophagy in intestinal epithelial cells ...

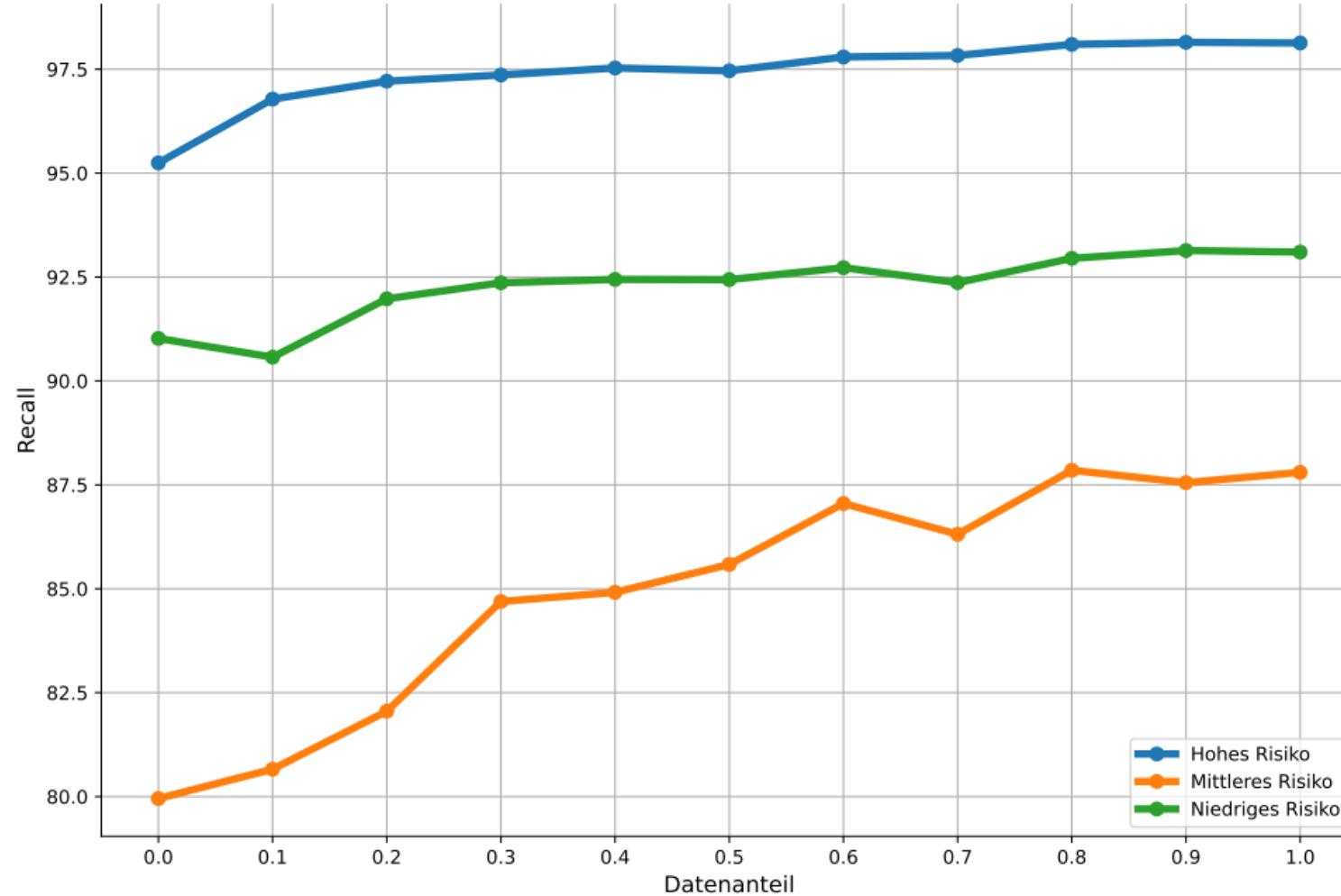
# Wie viele Trainingsdaten sind erforderlich?

Lernkurven für alle Textstellen aus dem **OLG-Goldstandard**



# Wie viele Trainingsdaten sind erforderlich?

Lernkurven nach Risikoniveau (OLG-Goldstandards)



# Evaluation der Domänenanpassung auf AG

Verringert die Domänenanpassung die Qualität des ursprünglichen AG-Modells?

Rechtsgebiet	Precision	Recall	Continual	Recall (PII)	Continual (PII)
Mietrecht	96.20%	95.74%	94.61%	98.90%	99.23%
Verkehrsrecht	97.23%	97.46%	95.38%	99.34%	99.34%
Gesamtergebnis	96.77%	96.73%	95.04%	99.16%	99.29%

## 4. Fazit & Ausblick

- Automatische Anonymisierung deutscher Gerichtsentscheidungen ist **ein gelöstes Problem**
- Allerdings benötigen wir einen großen und hochwertigen Goldstandard, um ein verlässliches Modell zu trainieren
- Sehr gute Ergebnisse können bereits mit dem deutschen GottBERT-Modell (125 M Parameter) erzielt werden
- Aber nur **eingeschränkte Robustheit** der hochspezialisierten Modelle
- Modelle bereits anwendbar für einige andere Rechtsgebiete und Instanzen (zumindest bei PII)
- Allerdings ist der Recall oft noch nicht ausreichend für eine vollautomatische Anonymisierung
- Domänenanpassung mit **in-domain-Trainingsdaten** scheint unabdingbar
- Erhebliche Mengen von Trainingsdaten v.a. für Pseudoidentifikatoren nötig, bessere Lernkurven bei PII
- Keine ersichtliche Verbesserung bei domänenübergreifende Robustheit durch den Einsatz von Data Augmentation
- Continual Learning erweist sich als gute Methode für eine robuste Domänenanpassung
- **Übertragung des Continual Learning-Ansatzes auf neue Textsorten?**

① Hilfe

## Risiko ②

Wie gefährlich die Veröffentlichung der Textstelle wäre.

Niedrig  Mittel  Hoch

## Mindestkonfidenz 0% ②

Je höher der Wert, desto sicherer ist die automatische Erkennung.

## Tags ②

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
- Natürliche Person • Identifizierendes Merkmal
- Juristische Person • Name j
- Adresse • Ortsangabe a
- Adresse • Identifizierendes Merkmal
- Fahrzeug • Identifizierendes Merkmal
- Datum • Prozessgeschichte d
- Datum • Sachverhalt s
- Datum • Weltwissen

## 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) -

Prozessbevollmächtigte zu 1), 2):

Rechtsanwälte Heinrich &amp; Kollegen, Züricher Straße 10, 90431 Nürnberg

gegen

Thomas Schütz, Feldstraße 4 d, 91096 Möhrendorf

- Beklagter zu 1) -

Luise Schütz, Feldstraße 4 d, 91096 Möhrendorf

- Beklagte zu 2) -

Prozessbevollmächtigte zu 1), 2):

Steinbrecher + Amberger Rechtsanwälte PartGmbB, Gohestraße 25, 91054 Erlangen

wegen Räumung

## Ausgewählter Abschnitt

Mozartstraße

 annotieren

## Ausgewählte Annotationen

Mozartstraße 23, 91052 Erlangen

   Adresse • Ortsangabe Akzeptiert  Nicht beibehalten alt + b Niedriges Risiko alt + r 99%

**Demonstrator**  
**AnGer 2025**

[corpora.linguistik.uni-erlangen.de/leak/](http://corpora.linguistik.uni-erlangen.de/leak/)

Amtsgericht Ingolstadt

Az: 13 C 510/20

Haakonweg 61, 85053 Ingolstadt

Telefon: 52838 / 528 - 74

Telefax: 42557 / 137 - 487

Verkündet am: 20. 8. 2020

(Schneider), JAng.

Urkundsbeamtin d. Geschäftsst.

IM NAMEN DES VOLKES

In dem Rechtsstreit

Xaver Heuer, Tuplenstraße 54, 85053 Ingolstadt

- Klägerin zu 1) -

Elodie Heuer, Tuplenstraße 54, 85053 Ingolstadt

- Kläger zu 2) -

Prozessbevollmächtigte zu 1), 2):

Rechtsanwälte Hanke &amp; Kollegen, Pilatusweg 32, 90491 Nürnberg

gegen

Hellmut Johann, Seeuferstraße 35e, 85132 Schernfeld

- Beklagter zu 1) -

Thea Johann, Seeuferstraße 35e, 85132 Schernfeld

- Beklagte zu 2) -

Realistisch

Abkürzen

Schwärzen

 Nur Fehlstellen anzeigen Automatische Mehrfachaktualisierung

✓ Amtsgericht Erlangen

 Formales • Gericht

Amtsgericht Ingolstadt

Amtsgericht Ingolstadt

✓ 11 C 122/20

 Formales • Aktenzeichen

13 C 510/20

✓ Mozartstraße 23, 91052 Erlangen

 Adresse • Ortsangabe

Haakonweg 61, 85053 Ingolstadt

✓ 09131/782-01

 Formales • Aktenzeichen

52838 / 528 - 74

✓ 09131/782-105

 Formales • Aktenzeichen

42557 / 137 - 487

✓ 23.7.2020

 Datum • Prozessgeschichte

20. 8. 2020

✓ Patrick Müller

 Natürliche Person • Name

Xaver Heuer

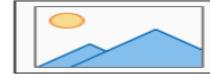
# Demonstrator

## AnGer 2025

[corpora.linguistik.uni-erlangen.de/leak/](http://corpora.linguistik.uni-erlangen.de/leak/)

# Auch andere Textsorten als Urteile?

Beispiele für Anwendung des AnGerProtoyps (ohne Domänenanpassung)



Das ist die Motivation hinter Hersheys neuer Partnerschaft mit der amerikanischen Behörde für Entwicklungszusammenarbeit (USAID) und ECOM, unserem größten Kakaolieferanten in Ghana. Letztes Jahr riefen wir ein kleines Pilotprogramm ins Leben, um Kleinbauern dabei zu helfen, ihre Kakaoproduktion zu steigern, die durch den Kakaobau bedingte Abholzung zu beenden und die Widerstandskraft der Kakaobäume zu verbessern. Die Initiative richtet ihr Augenmerk insbesondere auf zwei Herausforderungen, mit denen jeder Anbaubetrieb in Westafrika konfrontiert ist: Grundbesitz und Finanzierung. Laut Angaben der Lands Commission in Ghana haben weniger als 2 Prozent der 800.000 Kakaobauern des Landes einen rechtlichen Anspruch auf den von ihnen bewirtschafteten Grund und Boden. Vielmehr erhalten sie durch informelle Absprachen mit lokalen Anführern oder Grundbesitzern Zugang zu landwirtschaftlichen Flächen. Traditionellerweise ermöglichen diese mündlichen Vereinbarungen den Bauern die Rodung von Wäldern und die landwirtschaftliche Nutzung der entstandenen Flächen. Wenn die Kakaobäume allerdings nach etwa 30 Jahren – oder im Falle von Krankheiten früher – keinen Ertrag mehr abwerfen, müssen die Bauern vom ursprünglichen Grundbesitzer eine Genehmigung zur Wiederbepflanzung einholen. In Zeiten einer historisch hohen Nachfrage nach Grund und Boden weigern sich lokale Anführer und Landbesitzer zunehmend, dem Wunsch der Bauern nach Wiederbepflanzung zu entsprechen. Den Bauern bleiben zwei Möglichkeiten, von denen keine wünschenswert ist: entweder Urwälder roden und von vorne beginnen oder sich gänzlich aus dem Geschäft verabschieden. Die Partnerschaft mit USAID und ECOM zielt darauf ab, dieses Problem durch die Beseitigung mancher Hürden für die Wiederbepflanzung zu lösen. ECOM entwickelte ein innovatives Finanzierungsmodell, das den Bauern hilft, alte oder kranke Bäume zu entfernen und sie durch widerstandsfähige und ertragreichere Hybridsorten zu

**Risiko**  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.

- Sonstiges + Sonstiges
- Natürliche Person + Identifizierendes Merkmal
- Juristische Person + Name
- Adresse + Ortsangabe
- Adresse + Identifizierendes Merkmal
- Datum + Sachverhalt

German online news

Bisher galt es als böse Verschwörungstheorie und „Corona-Ketzerei“, wenn jemand behauptete, wir hätten es bei Corona mit einer „Test-Pandemie“ zu tun – und ohne die ausgiebigen Tests wäre die Situation nie so eskaliert und es gäbe keine Corona-Maßnahmen, wie wir sie kennen. Wer diese Position vertritt, muss sich auf Ausgrenzung, Diffamierung und Hass einstellen. Und das, obwohl sie kein Geringerer als das österreichische Gegenstück zum deutschen RKI-Chef Lothar Wieler vertritt: „Ohne PCR-Tests wäre die Pandemie niemandem aufgefallen“, sagte Professor Dr. Franz Allerberger, der Leiter der österreichischen Agentur für Gesundheit und Ernährungssicherheit (AGES). Ich habe hier bereits im Juni darüber berichtet. In den großen Medien wurde die Aussage aus berufenem Munde allerdings so gut wie totgeschwiegen. Wie so vieles, was nicht in das offizielle Corona-Narrativ passt. Und nun das! Kein Geringerer als Jens Spahn, seines Zeichens Gesundheitsminister und einer der obersten Betreiber der Corona-Panik, sagt kaum verklausuliert nichts anderes als der österreichische Wieler. Und wieder hören die meisten weg, und man tut so, als sei die Aussage gar nicht gefallen.

**Risiko**  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.

- Sonstiges + Sonstiges
- Formales + Aktenzeichen
- Natürliche Person + Name
- Natürliche Person + Identifizierendes Merkmal
- Juristische Person + Name
- Adresse + Ortsangabe
- Datum + Prozessgeschichte
- Datum + Sachverhalt

German Telegra

Demonstrator  
AnGer 2025

[corpora.linguistik.uni-erlangen.de/leak/](http://corpora.linguistik.uni-erlangen.de/leak/)

---

**Vielen Dank  
für Ihre Aufmerksamkeit!**