

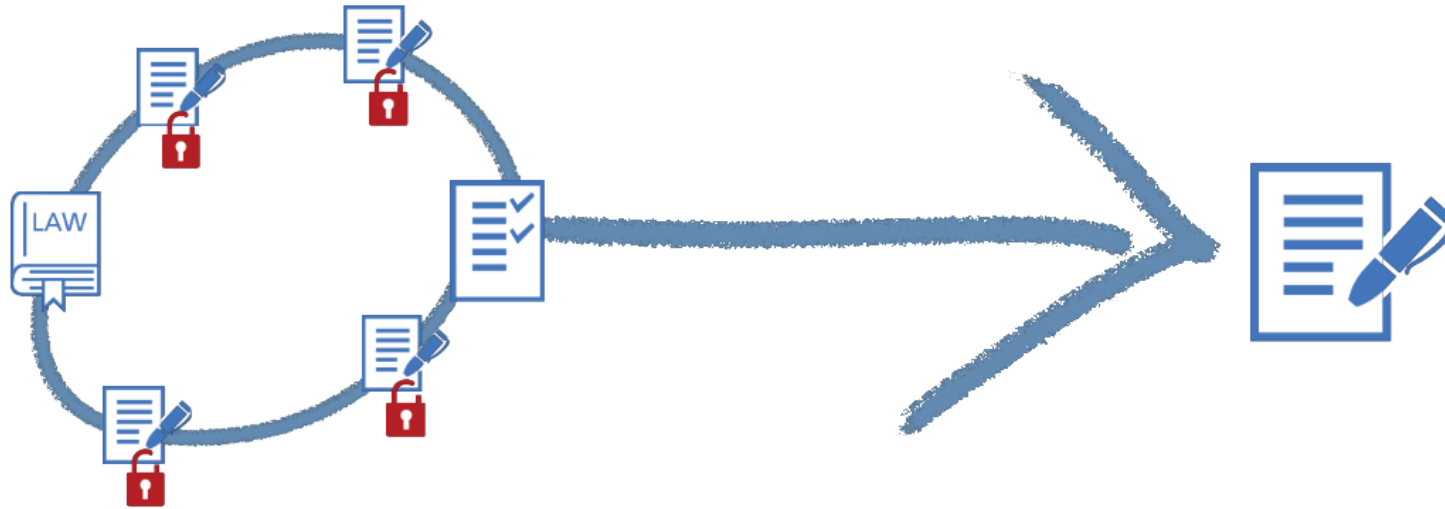


Combining Rule-Based and Machine Learning Methods for Efficient Information Extraction from Enforcement Decisions

Harry Nan (TiU / UvA), Maarten Marx (UvA), Johan Wolswinkel (TiU)

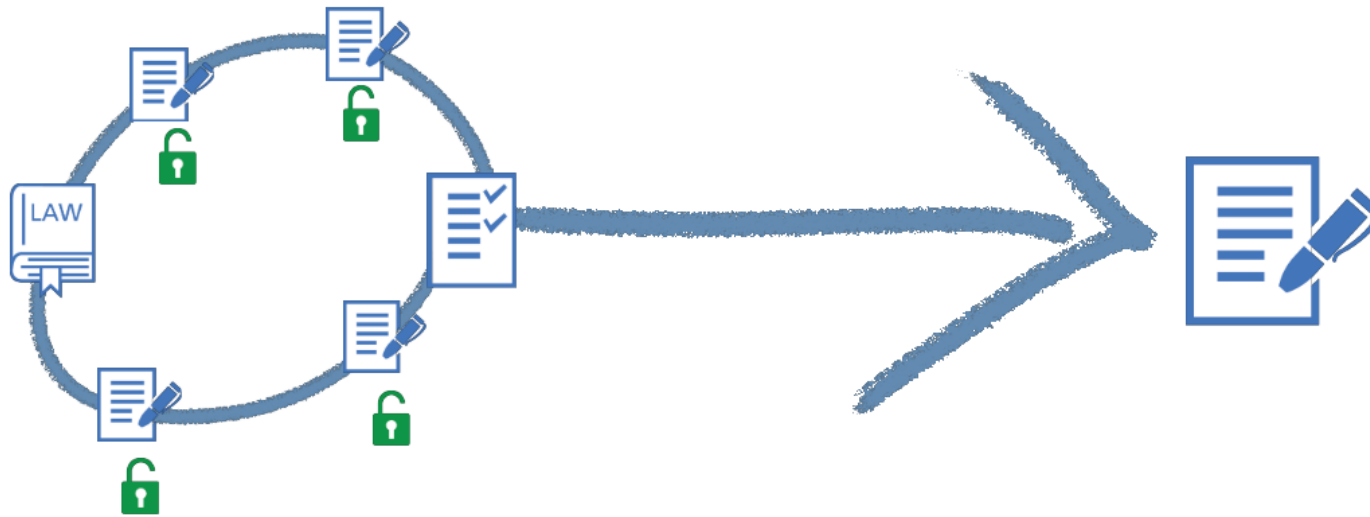
Background: open government

- Legal (AI) research tends to focus on statutes and court decisions
 - Why not administrative decisions as 'sources of law'?
 - Lower place in hierarchy
 - Less legalese
 - Usually not public



Background: open government

- But what if administrative decisions would be public?



Background: open government in the Netherlands

- Administrative decisions *shall* be public
 - As of 2026 (expected)
 - In machine-readable format
 - On a centralized portal

Staatsblad
van het Koninkrijk der Nederlanden



Jaargang 2021

499

Wet van 25 oktober 2021, houdende regels over de toegankelijkheid van informatie van publiek belang (Wet open overheid)

Until then:

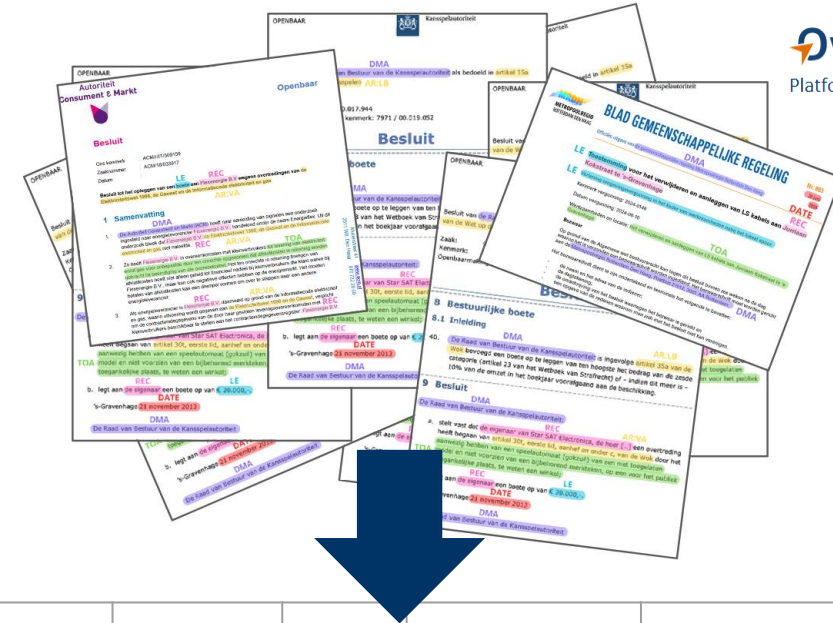
- Decentralised disclosure (if any)
 - No uniform template
- Wide variety
 - Licensing, Subsidizing, Enforcing...
- Poor data quality
 - Lacking metadata



Research aim: accurate metadata extraction

- Metadata administrative decisions
 - Date
 - Competent authority (DMA)
 - Recipient
 - Legal basis
 - Legal effect (e.g. monetary sum)
 - ... [dependent on type of decision]

- Example: enforcement decision
 - Legal provision violated (Violated Article)
 - Misbehaviour



Date	Recipient	Fine	Violated Article	Motivation
04/12/2013	De heer [...]	20000	artikel 30h, eerste lid van de Wok	Het exploiteren van vijf, respectievelijk zeven kansspelautomaten op een voor het publiek toegankelijke plaats
23/03/2023	Imperial E-Club Limited	100000	artikel 1, eerste lid, aanhef en onder a, van de Wok	Online kansspelen aangeboden zonder vergunning op de Nederlandse markt
15/09/2020	Probe Investments Limited	15000	artikel 1, eerste lid, aanhef en onder a, van de Wok	Het geven van gelegenheid om mede te dingen naar prijzen of premies zonder de vereiste vergunning

Method: data collection

- Enforcement decisions

- Administrative fines
- Administrative penalties
- ...

- Administrative bodies

#1. *Kansspelautoriteit* (Dutch Gambling Authority)

#2. *Autoriteit Financiële Markten* (Dutch Financial Markets Authority)

Enforcement decisions (299)	#1: KSA (110)	#2: AFM (179)
Administrative fines (175)	65	110
Administrative penalties (124)	55	69

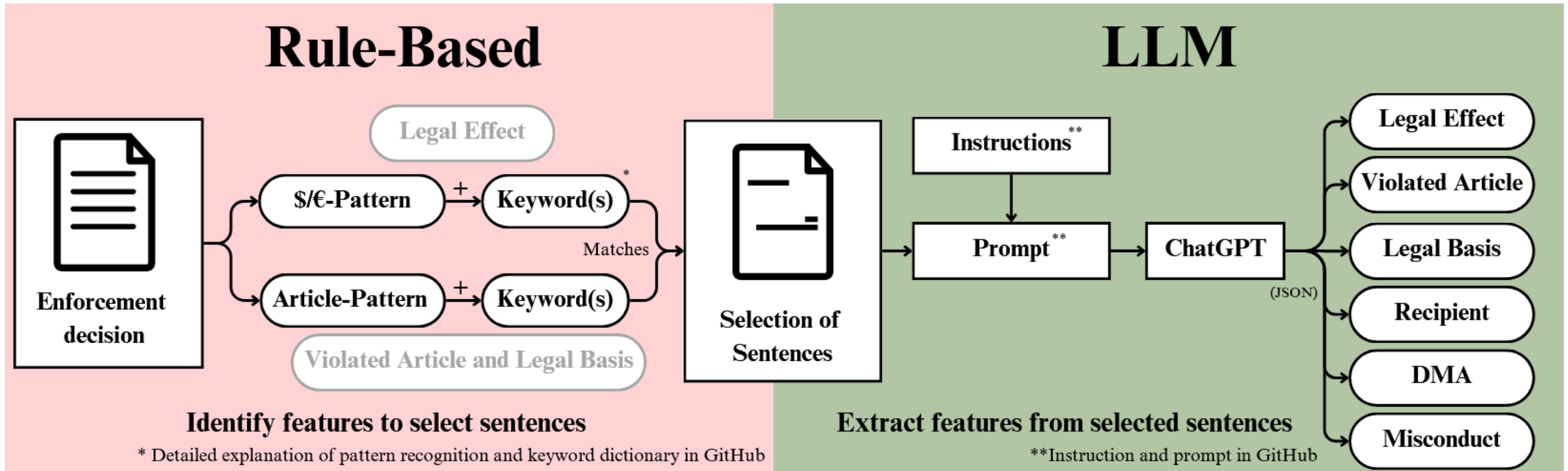
Method: two-step hybrid approach

- Combination of rule-based with LLM (GPT-3.5-turbo)

De Kansspelautoriteit heeft vastgesteld dat Clipboard Publications B.V. reclame maakt voor kansspelen waarvoor geen vergunning op grond van de Wet op de Kansspelen (hierna: Wok) is verleend. Daarmee handelt Clipboard Publications B.V. in strijd met de **Misconduct** van de Kansspelautoriteit (hierna: de raad van bestuur) van Clipboard Publications B.V. op het maken van reclame voor kansspelen waarvoor geen vergunning ingevolge de Wok is verleend (overtreding van artikel 1, eerste lid, onder b, van de Wok), te staken en gestaakt te houden door middel van het opleggen van een last onder dwangsom. **Violated Article**

De Autoriteit Financiële Markten (AFM) heeft besloten aan Esperite **Legal Effect** bestuurlijke boete van €312.500 op te leggen, omdat Esperite niet heeft voldaan aan de op haar als uitgevende instelling rustende verplichting om haar financiële verslaggeving tijdig algemeen verkrijgbaar te stellen. Esperite heeft: **Misconduct** **Violated Article**

De Autoriteit Financiële Markten (AFM) heeft besloten aan u een bestuurlijke boete van €4.000 op te leggen wegens het feitelijk leiding geven aan overtreding van artikel 4:11, tweede lid, van de Wet op het financieel toezicht (Wft) door Perfect Hypotheken B.V. (Perfect). De boete wordt opgelegd, omdat Perfect in de periode van 14 juli 2009 tot en met 2 mei 2013 geen adequaat beleid heeft gevoerd dat een integere uitoefening van haar bedrijf waarborgde en u als enig (middellijk) bestuurder van Perfect feitelijk leiding heeft gegeven aan deze overtreding.



Nan et al. (2024).

Results

- Overall: pretty good performances
- But:
 - Legal basis (# 2)?
 - Penalty slightly worse performing than fine?
- Explanations:
 - Differences across administrative bodies
 - Differences in types of administrative decisions

Table 1. Precision, Recall, and F1-scores for two governmental bodies for six identified information types [1].

Category	Feature	Administrative Body					
		# 1			# 2		
		Precision	Recall	F1-score	Precision	Recall	F1-score
Fines	Legal Effect	0.950	1.000	0.974	1.000	1.000	1.000
	Violated Article	1.000	1.000	1.000	1.000	0.900	0.947
	Legal Basis	1.000	1.000	1.000	0.222	0.222	0.222
Non-Identifiable features	Recipient	0.800	0.833	0.816	1.000	1.000	1.000
	DMA	0.874	0.920	0.896	1.000	1.000	1.000
	Misconduct	0.811	0.818	0.814	1.000	1.000	1.000
Penalties	Legal Effect	0.750	0.789	0.769	0.625	0.676	0.650
	Violated Article	0.700	0.737	0.718	0.658	0.738	0.696
	Legal Basis	1.000	1.000	1.000	0.300	0.300	0.300
	Recipient	0.818	0.825	0.821	0.797	0.800	0.798
	DMA	0.801	0.840	0.820	0.900	0.900	0.900
	Misconduct	0.954	0.958	0.956	0.659	0.724	0.690

Nan et al. (2024).

Further research directions

- Future (I): information extraction
 - Generalizability to other types of administrative decisions and administrative bodies
 - Comparison with other (non-hybrid) methods for metadata extraction
 - Rule-based
 - LLMs
- Future (II): building on metadata
 - Measuring similarity of administrative decisions
 - Analysis / finding patterns in administrative decision-making



Thank you!

