



chai Justice

Natural language processing agency



Summary

Part 1 : Introduction

Project presentation
Project structure
Project team

Part 2 : Preparation

Data cleaning
Corpus analysis
Special outliers
Document segmentation
Trends in topics

Part 3 : Semantic analysis

Context of occurrences
Specific axis
Key words in context

Part 4 : Structural analysis

Grammatical roles
AMR framing of justice
AMR graphics examples
AMR Analysis of "Justice"

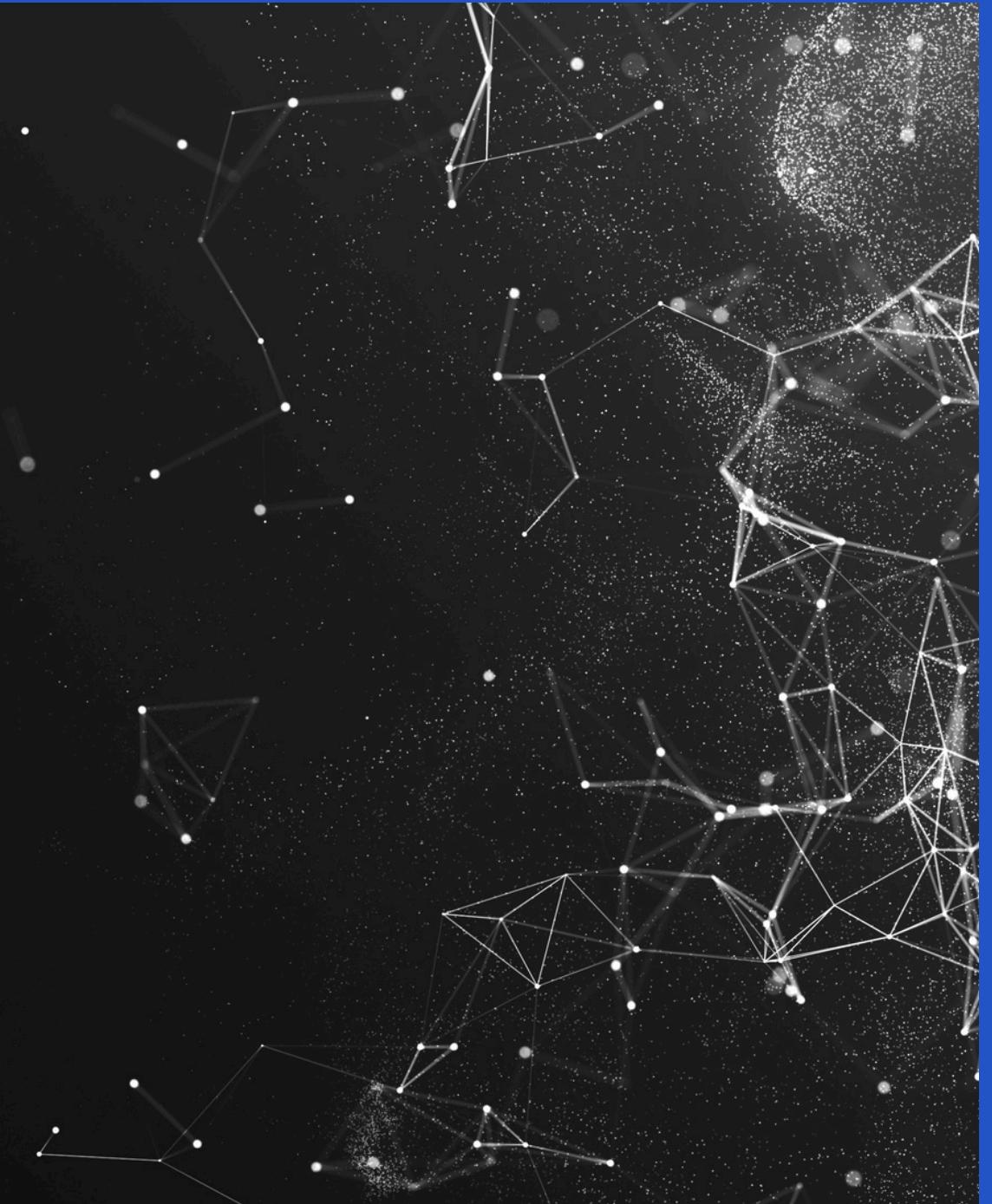
Part 1

INTRODUCTION

Project presentation

Project structure

Project team



Project presentation

Analysis of AI ethics charters through semantic and AMR graph exploration context

- Recent years have seen a surge of initiatives aiming to define what “ethical AI” means.
- Numerous charters and manifestos have been published by academia, industry, regulatory bodies, etc.
- These documents reflect diverse visions, values, and priorities related to AI ethics.

Project Objective

Studying the linguistic and semantic specificities of a corpus of AI ethics charters and manifestos, focusing on the term justice and its contextual meanings, through:

- Linguistic analysis of the corpus
- Semantic abstraction using AMR (Abstract Meaning Representation)

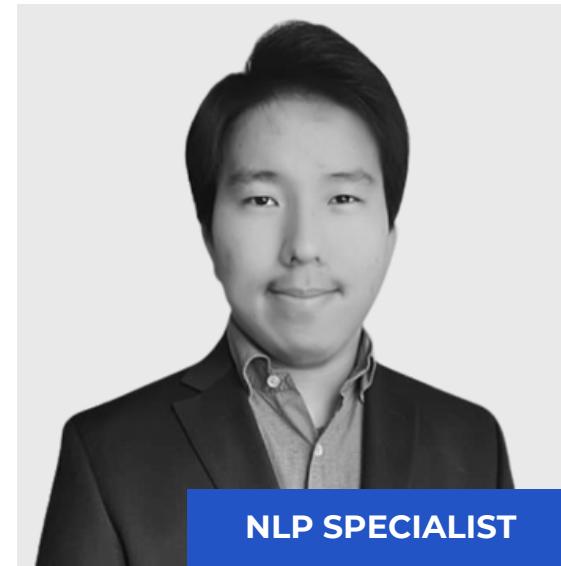
Our Team

Our team brings together diverse backgrounds in data science, linguistics, and artificial intelligence, combining technical expertise with a strong interest in ethical issues surrounding AI.

Each of us contributed to different aspects of the project — from data collection and analysis to semantic modeling and AMR graph interpretation.



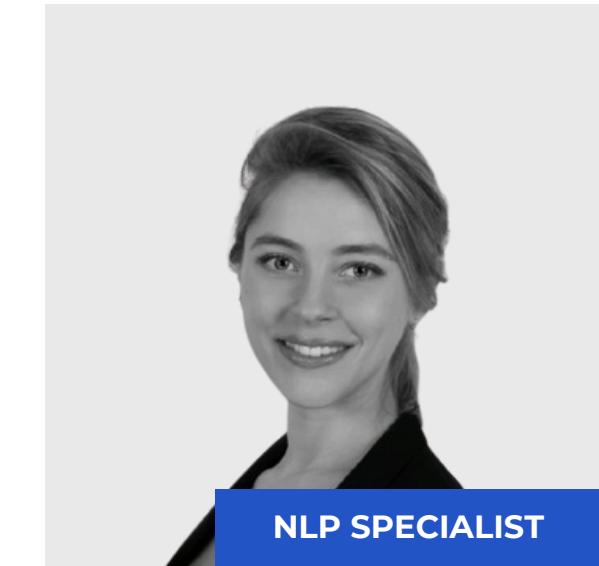
Paul Cintra



Yimou Zhang



Antoine Durand



Cora Bouvier

Project structure

3 steps

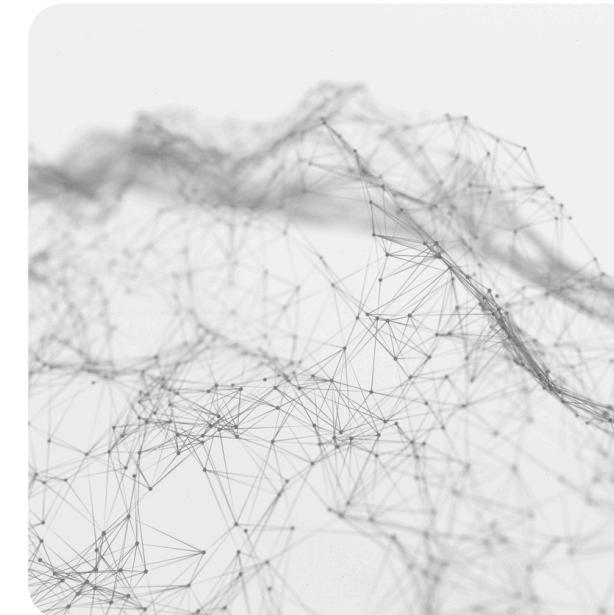
Each step explores a different level of analysis, from raw data exploration to semantic interpretation through AMR graphs.



Data retrieval & exploration



Exploration of the term justice



Justice in AMR graph representations

Reference : Mapping AI Ethics: a meso-scale analysis of its charters and manifestos — M. Gornet et al., 2024

Part 2

PREPARATION

Data cleaning

Corpus analysis

Special outliers

Document segmentation

Trends in topics



Data cleaning

Observations

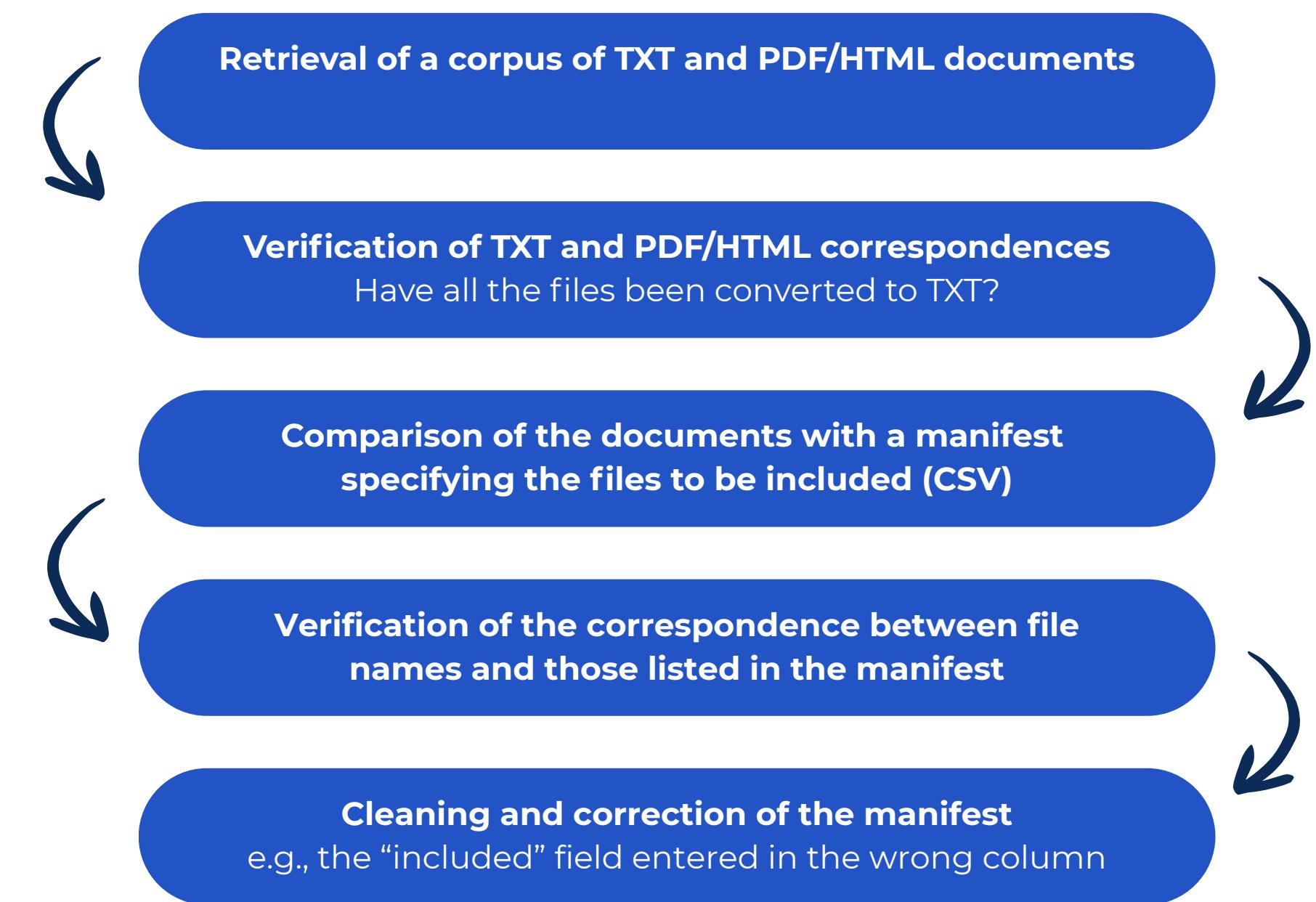
Number of PDF files: 801

Number of TXT files: 650

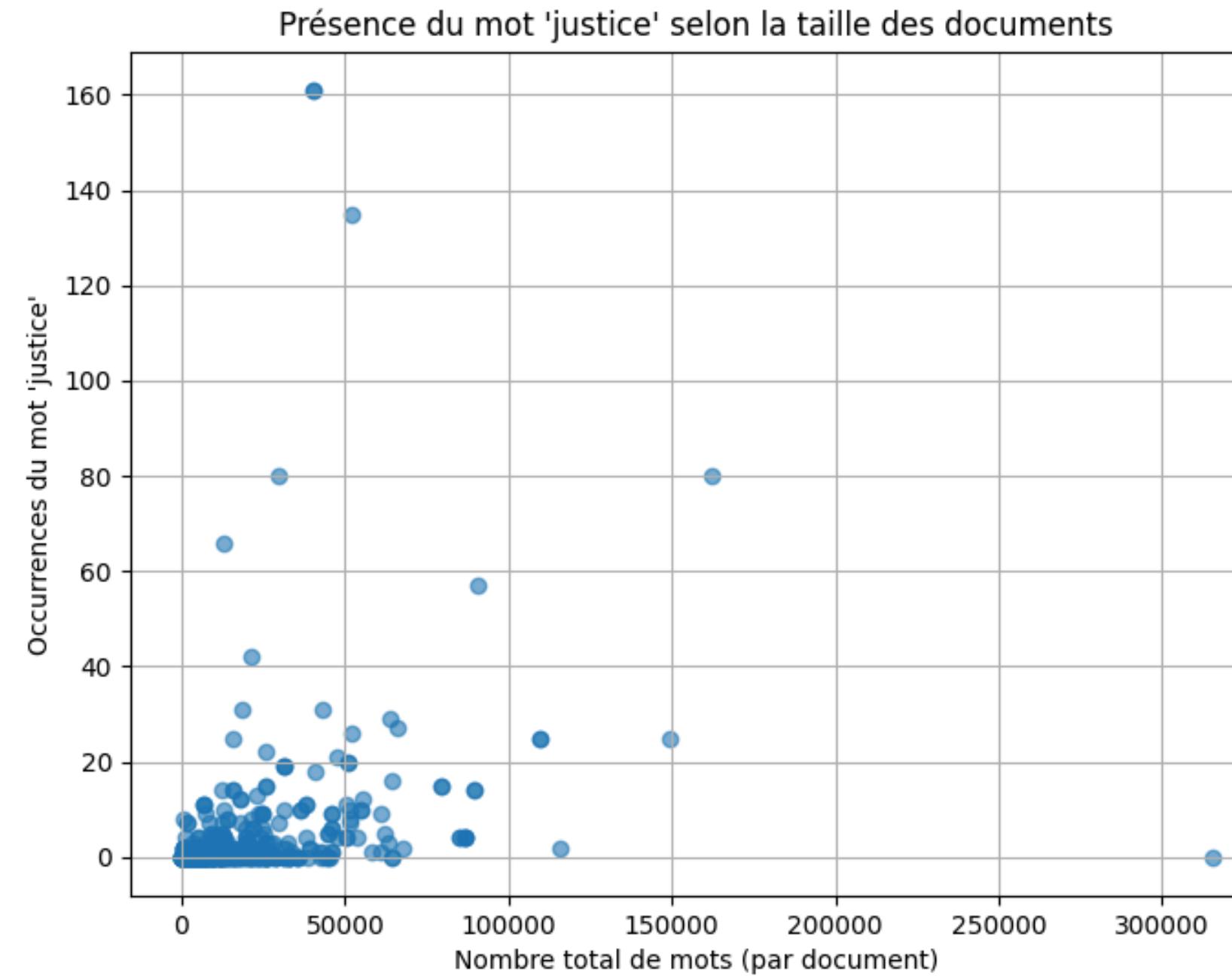
Total number of files to be included: 437

Number of missing files in TXT: 43

Number of missing files in PDF/HTML: 43



Corpus Analysis



Justice occurrence

Low frequency overall

Most documents contain very few or no mentions of the term justice, even when they are relatively long. Most points are concentrated below 50 occurrences.

Outliers identified

A few documents show a much higher frequency, with up to 200 mentions of justice, standing out as clear outliers in the dataset.

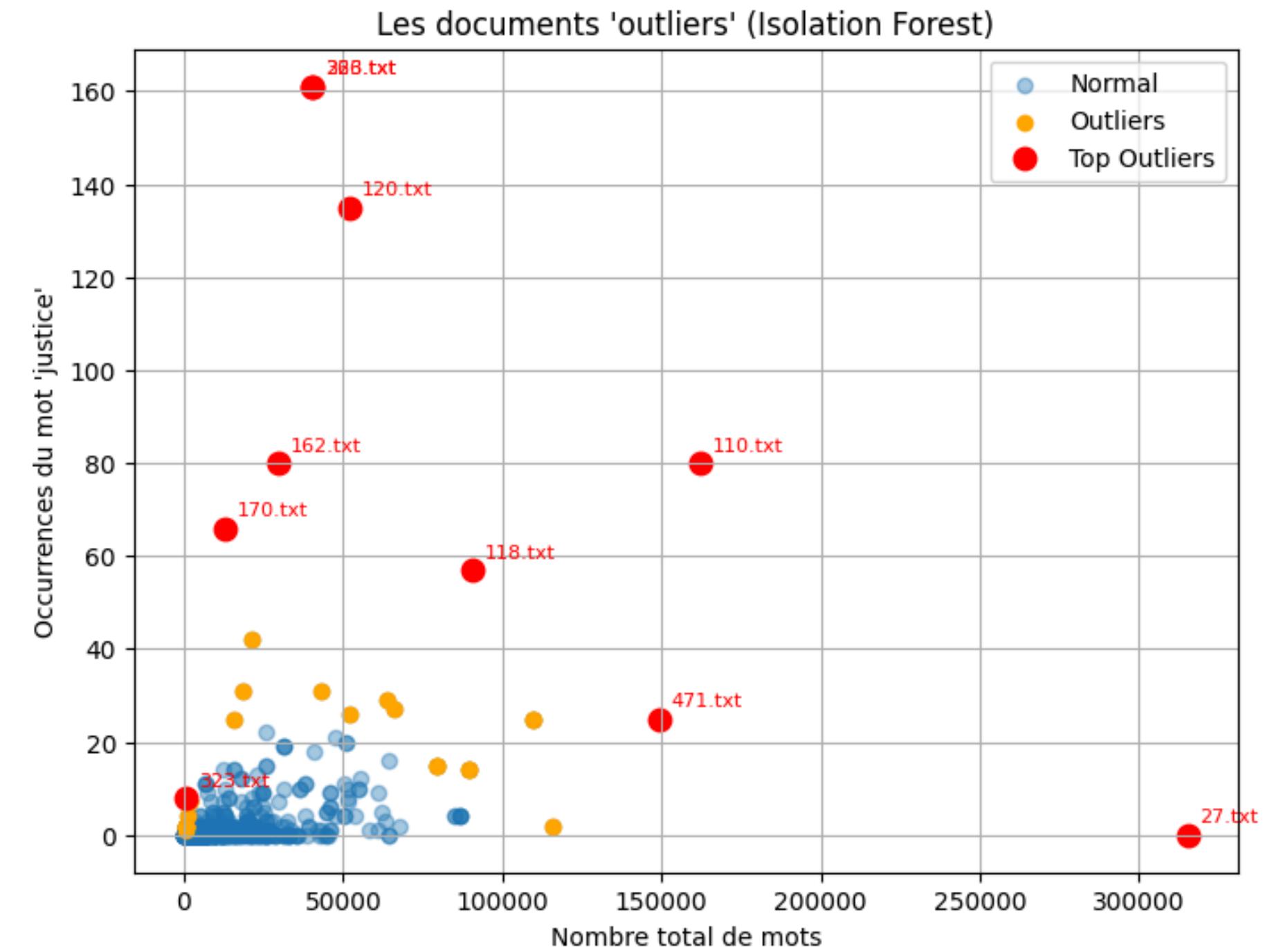
No proportional relation

There is no direct correlation between document length and number of times justice appears. Some long texts mention it rarely, while shorter ones use it a lot.

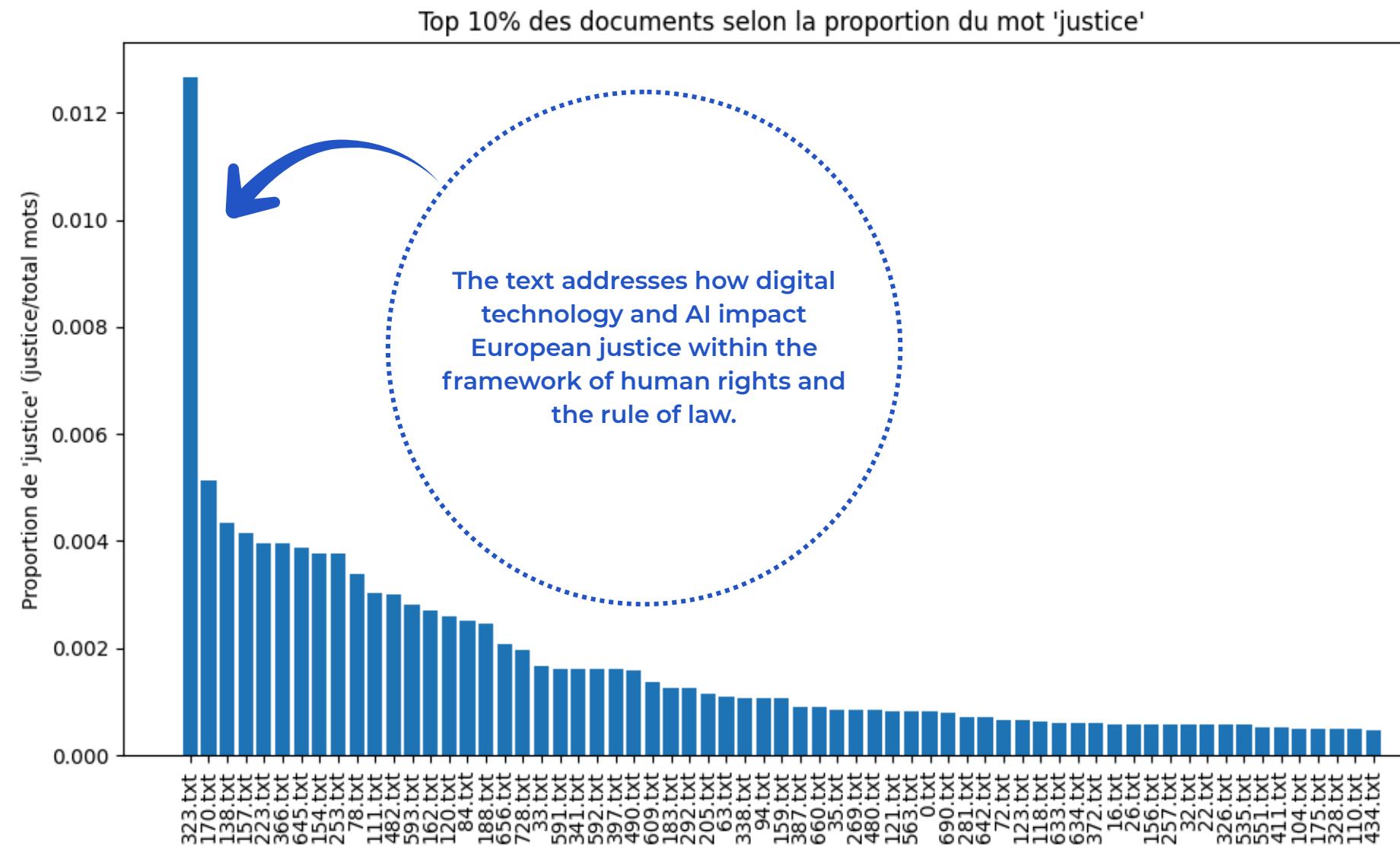
Special outliers

Exception

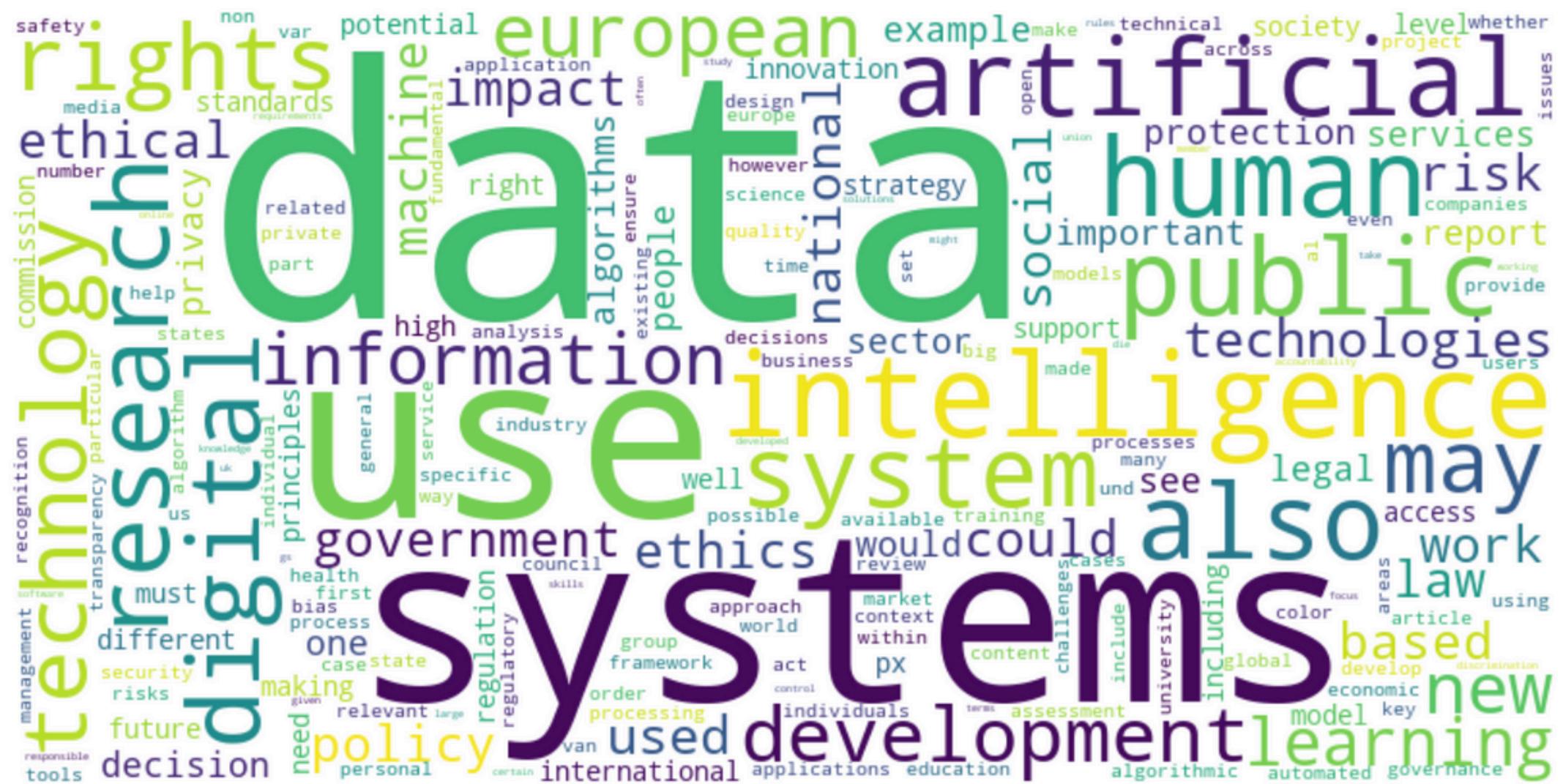
Some outliers in orange and top outliers in red deviate from the general trend. These deviations can result from either an unusually high number of occurrences of justice or an atypical document length, with some long texts mentioning the term very rarely or not at all. This suggests that justice is not proportionally distributed across documents, and its presence reflects specific thematic or stylistic choices rather than document size.



Document segmentation



Trends in topics



Technological axis

data, systems, algorithms, learning

Ethical and institutional axis

human rights, public policy, governance

- A predominance of terms such as **data, systems, use, intelligence, human, and public** reflects the central themes of AI ethics charters.
- The frequent occurrence of **data, systems, and use** highlights a strong focus on the technical and operational dimensions of artificial intelligence.
- Words like **human, rights, ethics, and public** indicate societal and normative concerns, centered on fundamental rights, accountability, and transparency.
- Finally, the presence of **policy, government, law, and research** shows that reflections on AI ethics are also grounded in institutional and regulatory frameworks.

Part 3

SEMANTIC ANALYSIS

Context of occurrences

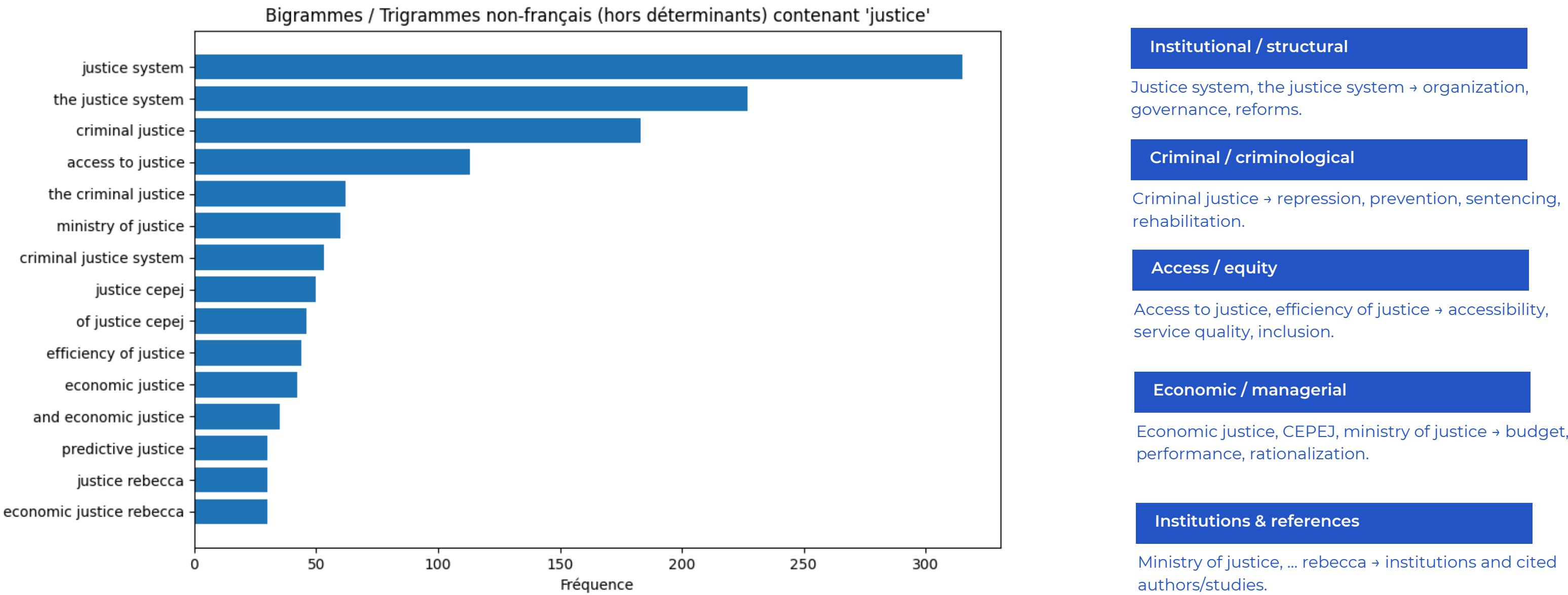
Specific axis

Key words in context

AMR Analysis of "Justice"

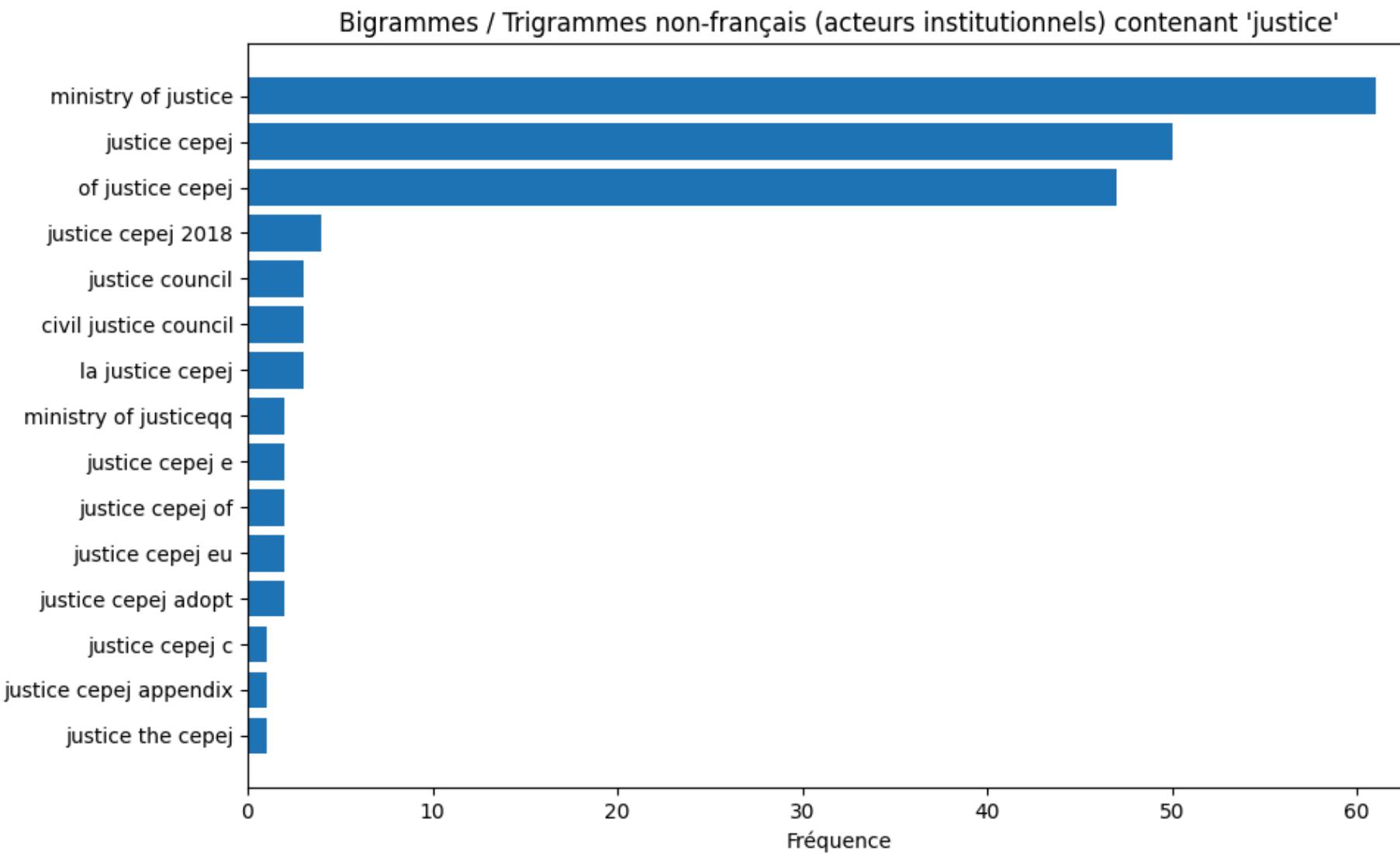


Context of occurrences



So what? Primary focus is system & governance. Other axes specify security, access, and performance.

Institutional structures

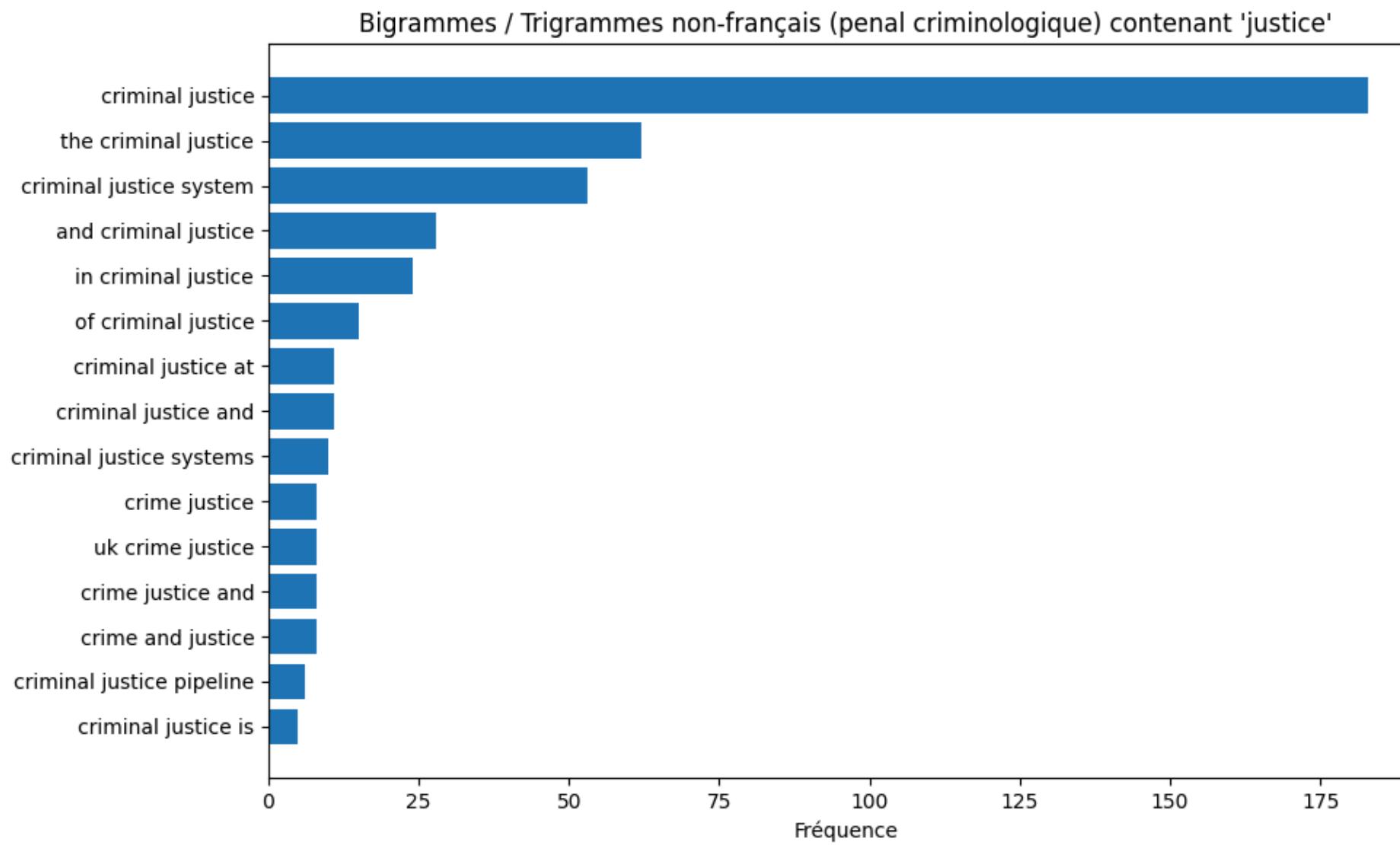


Institutional-structural: "Justice system" dominates, framing justice through organizational and administrative structures.

Policy framing: expressions like court of justice and administration of justice point to institutional governance and procedural focus.

Long tail: minor forms reflect variations in terminology across institutional or national contexts.

Penal & criminological

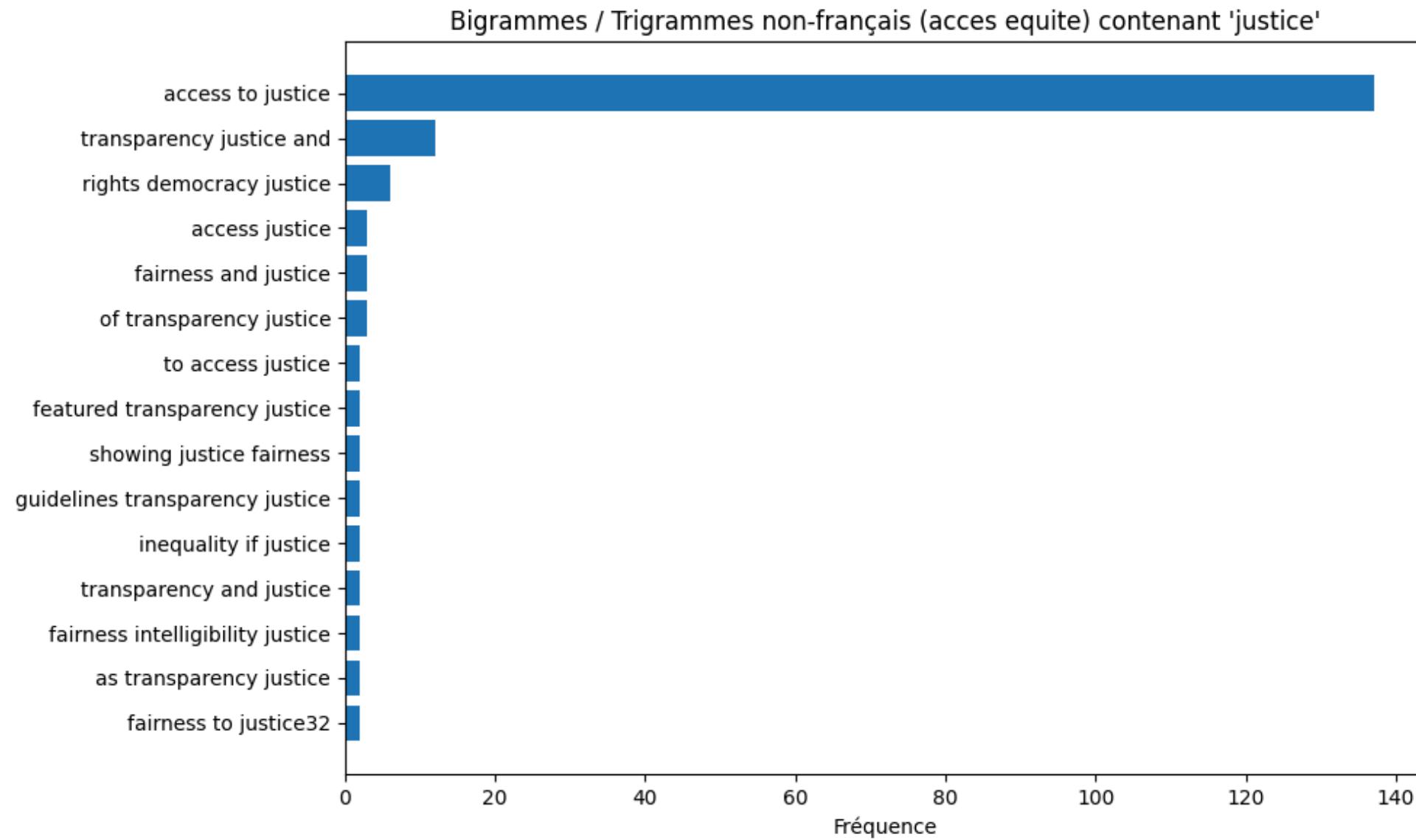


Justice-centric: "Criminal justice" overwhelmingly dominates, framing justice through institutional and system-oriented language.

System framing: recurrent forms like criminal justice system or the criminal justice highlight structural and policy dimensions.

Long tail: variants such as crime justice or crime and justice suggest peripheral, thematic, or localized usages.

Equity and accessibility

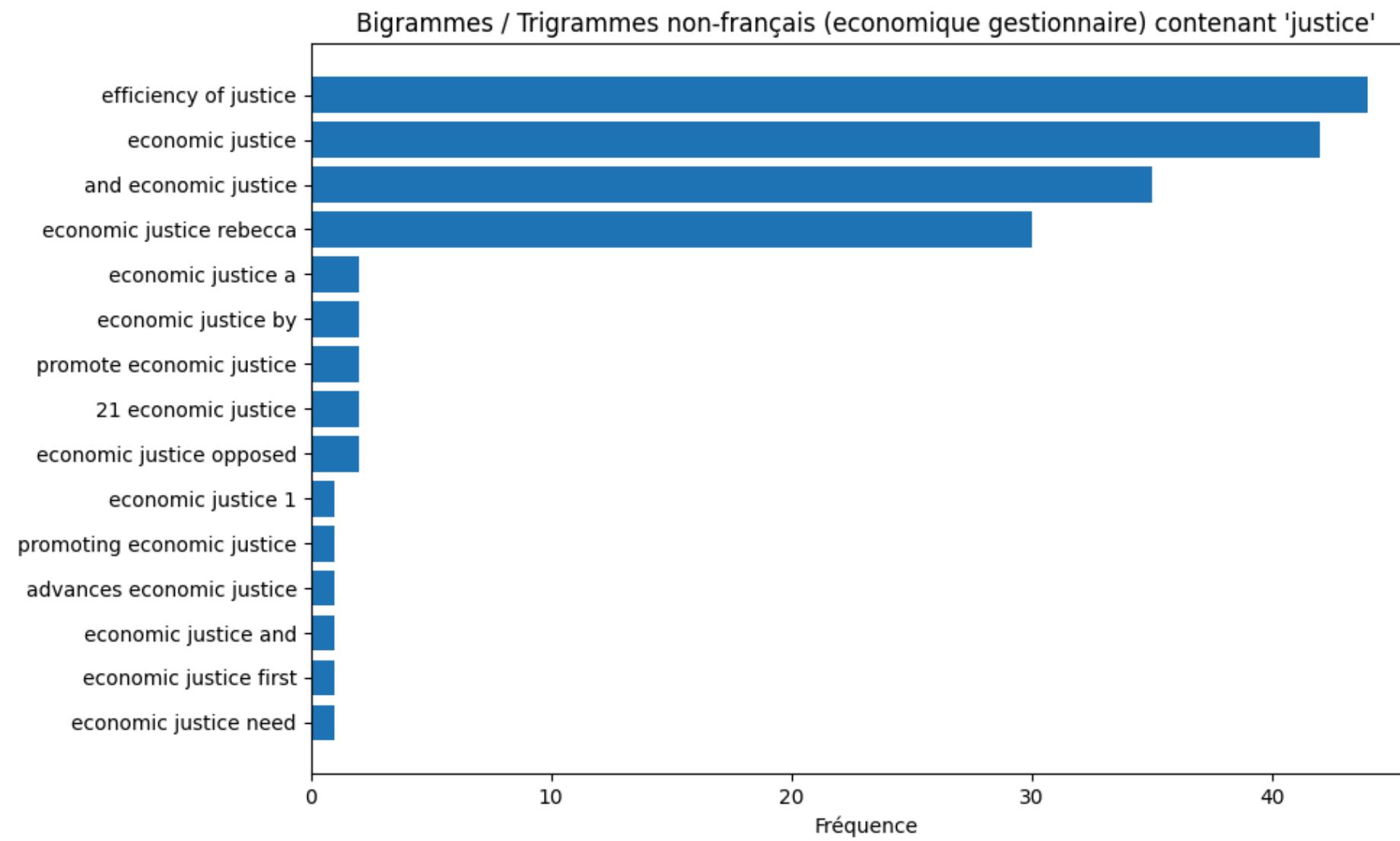


Equity-centric: "Access to justice" dominates overwhelmingly, anchoring justice within inclusion, fairness, and rights discourse.

Transparency framing: co-occurrences with transparency and democracy indicate a governance and accountability perspective.

Long tail: scattered terms such as fairness and justice or guidelines transparency justice suggest normative or procedural interpretations of equitable access.

Economic efficiency

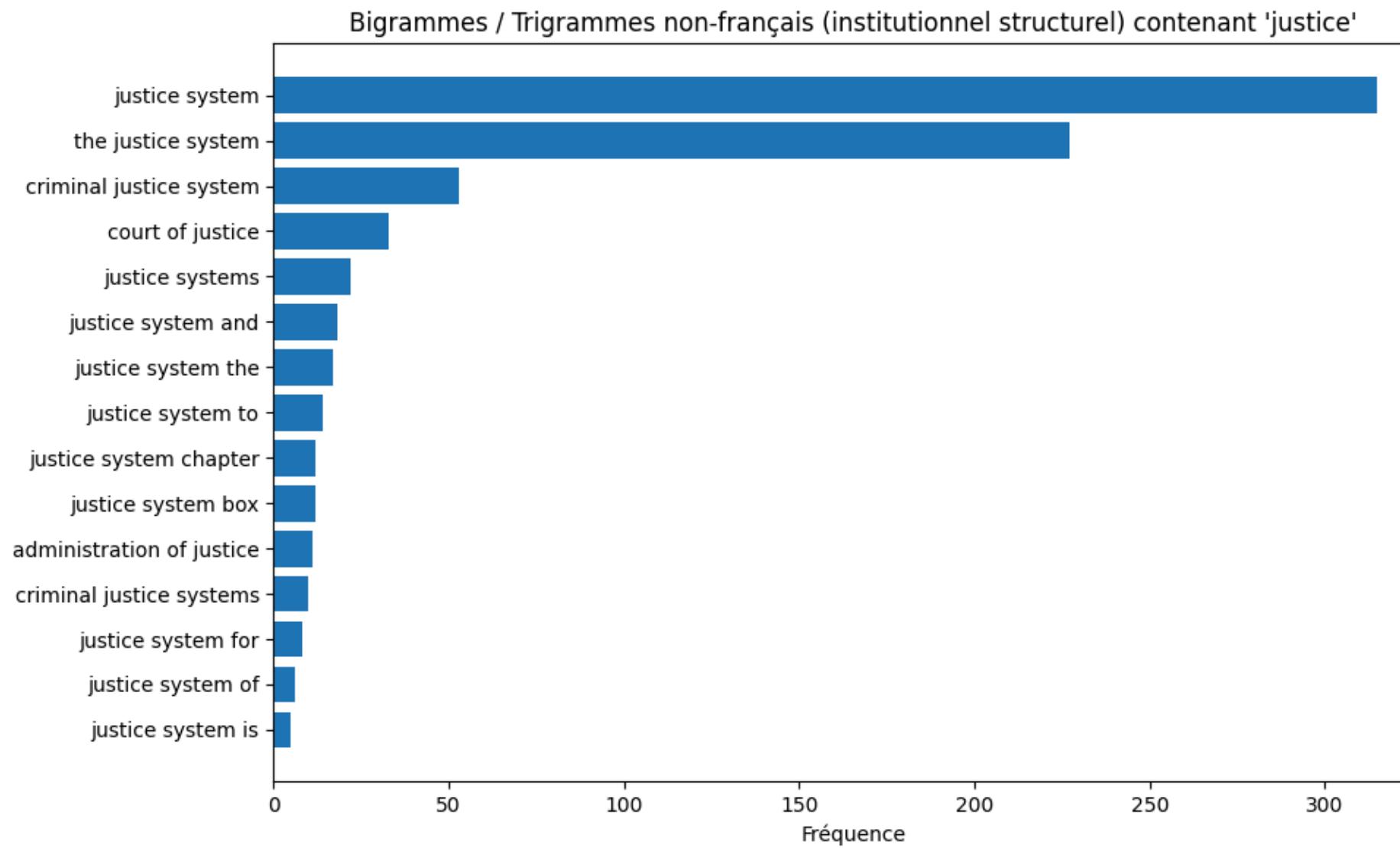


Economic-centric: “Efficiency of justice” and “economic justice” dominate, framing justice through performance and resource allocation.

Managerial framing: combinations like and economic justice or promoting economic justice point to policy and advocacy discourses linking justice to growth and inclusion.

Long tail: scattered terms such as advances economic justice or economic justice need suggest developmental and rights-based interpretations.

Institutional actors



- **Institution-centric :** Ministry of Justice dominates, CEPEJ is the main secondary actor → policy/benchmarking lens.
- **EU framing :** CEPEJ signals European governance & performance (metrics, reports, years like 2018).
- **Long tail :** councils indicate regulatory/advisory sources.



Key words in context

--- Exemples KWIC (Key Word In Context) ---

- Comment les acteur.rice.s de la [justice] perçoivent-ils l'avènement de cette technolo
- ting Citizenship and living in the UK Crime, [justice] and the law Disabled people Driving and tran
- of inference and correlation in the criminal [justice] system, and suggested that its use in the US
- OPINION OF THE COMMITTEE ON CIVIL LIBERTIES, [JUSTICE] AND HOME AFFAIRS
- democracy, the rule of law and distributive [justice] , or on the human mind itself .) Adopt adequa
- e utmost precaution , notably in the area of [justice] and law enforcement; 68.
- ally discriminatory practices, from criminal [justice] to education to employment, and presents ris
- democracy, the rule of law and distributive [justice] , or on the human mind itself.) Adopt adequat
- sioner, I aim to promote economic and social [justice] through consumer protection and competition
- ield of civil, commercial and administrative [justice] h and, their application in the field of civ

Voisins à gauche (± 3 mots, sans stopwords)

Part 4

STRUCTURAL ANALYSIS

Grammatical roles

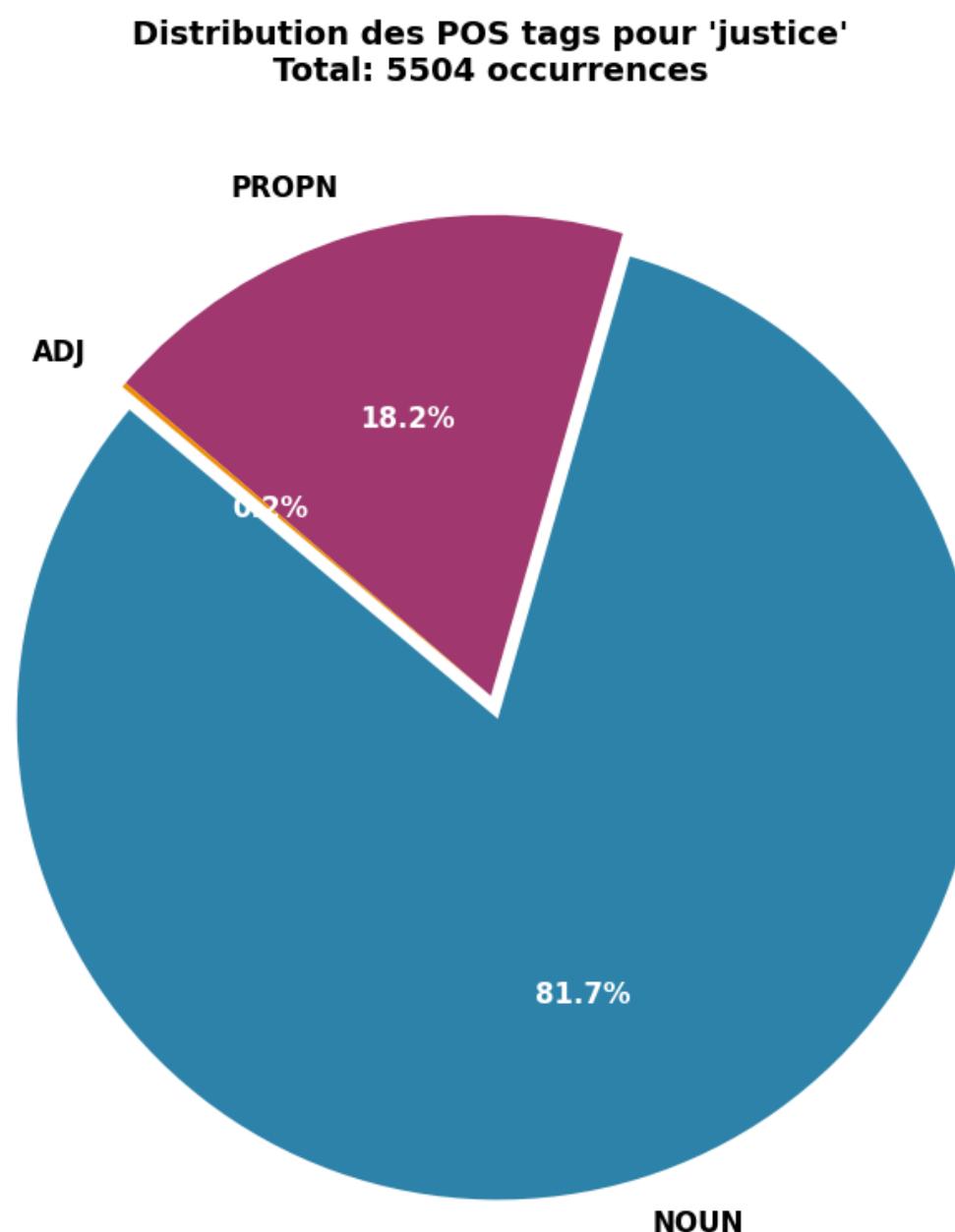
AMR framing of justice

ARM graphics examples

AMR Analysis of "Justice"



Grammatical roles



Syntactic Analysis of justice (POS tagging)

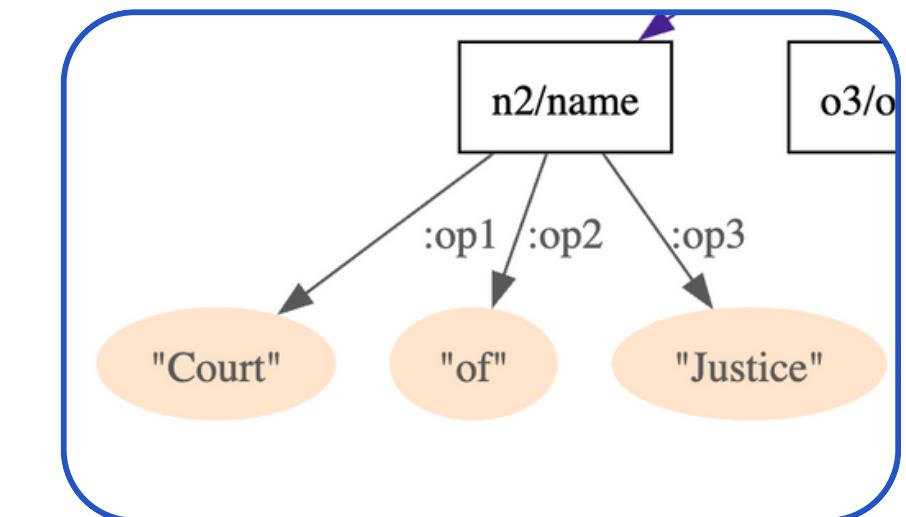
- **NOUN – 82%**: dominant use, referring to the concept of justice (access to justice, justice system).
- **PROPN – 18%**: institutional use (Ministry of Justice, Court of Justice).
- **ADJ – 0%**: annotation errors (e-Justice Strategy).

Insight: The term justice is mainly used as an abstract concept rather than an institutional reference, showing a theoretical and analytical orientation of the corpus.

AMR framing of justice

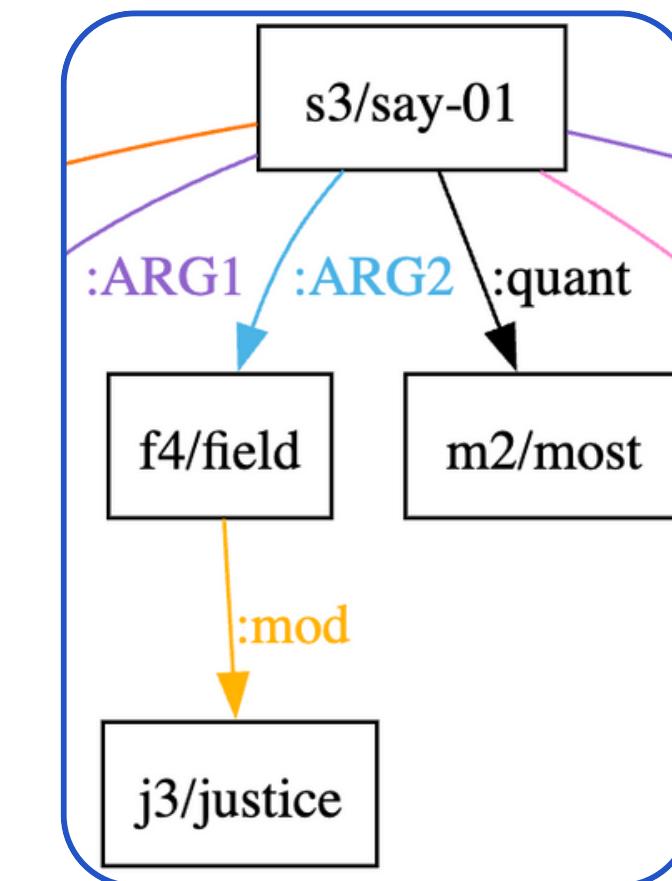
Concept level

Justice appears 346 times (after separating multi sentence) as a nominal AMR concept - **(j / justice)** - confirming its treatment as an abstract noun, not as an event or process.



PropBank alignment

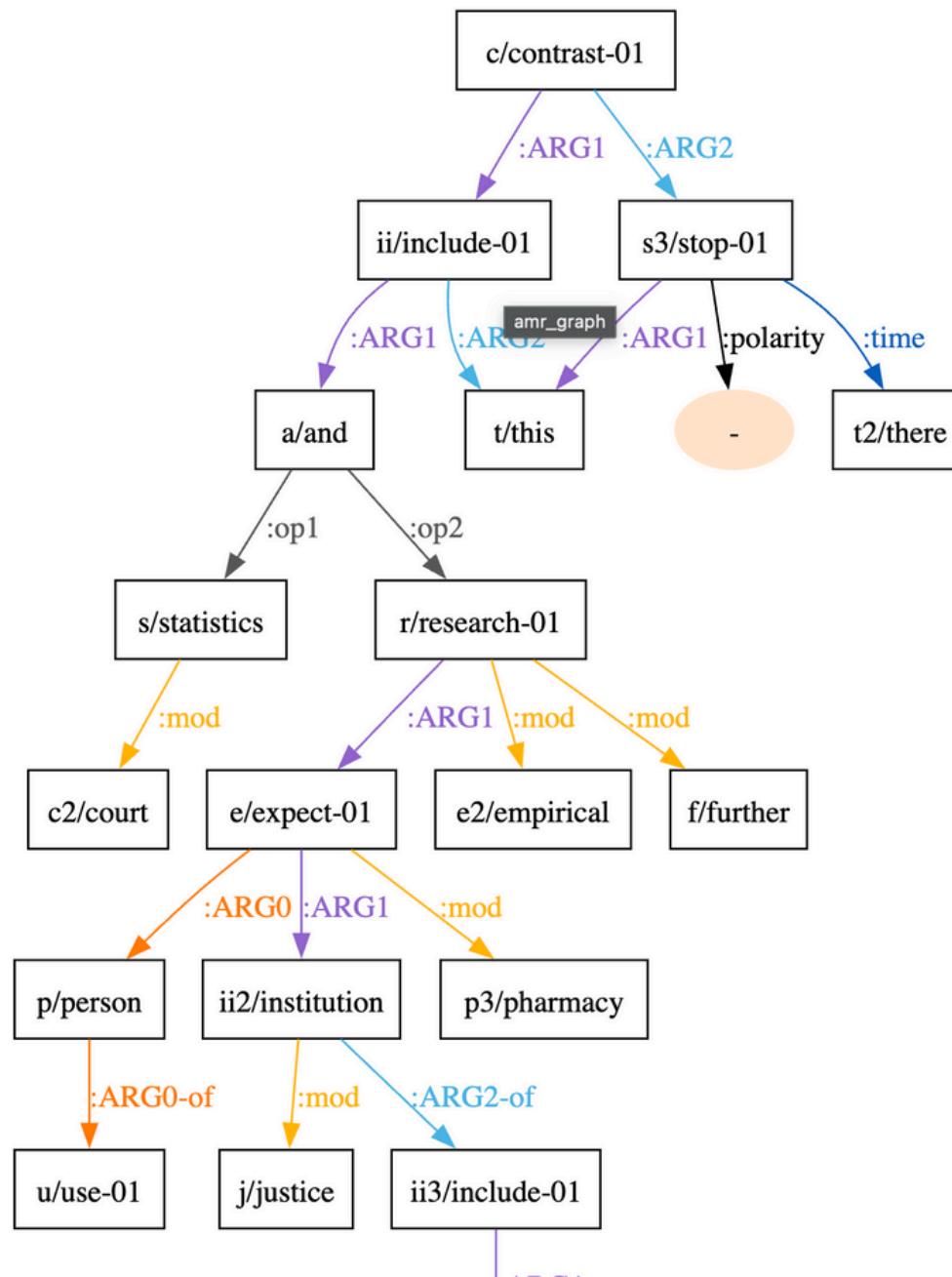
None — **PropBank only defines verb frames** (judge, justify), which link to judicial or evaluative actions rather than the concept of justice itself.



Relational framing

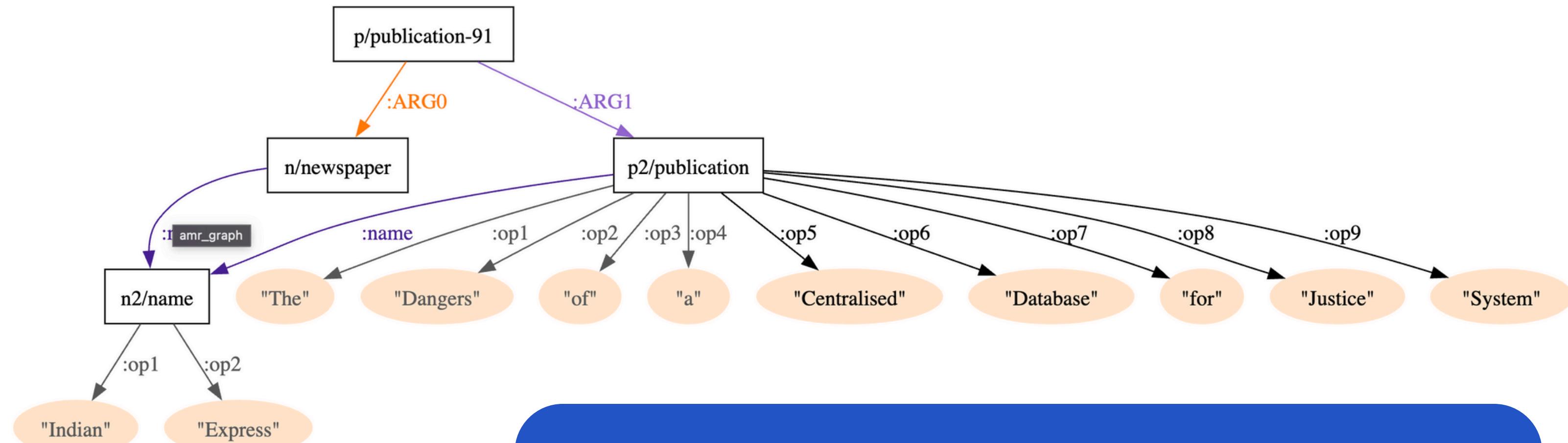
In AMR, justice is modified or contextualized through structures such as :
:mod → **social justice - modifier relation**
:name or :wiki → **Justice Department - named entity, not a concept node**

AMR graphics examples



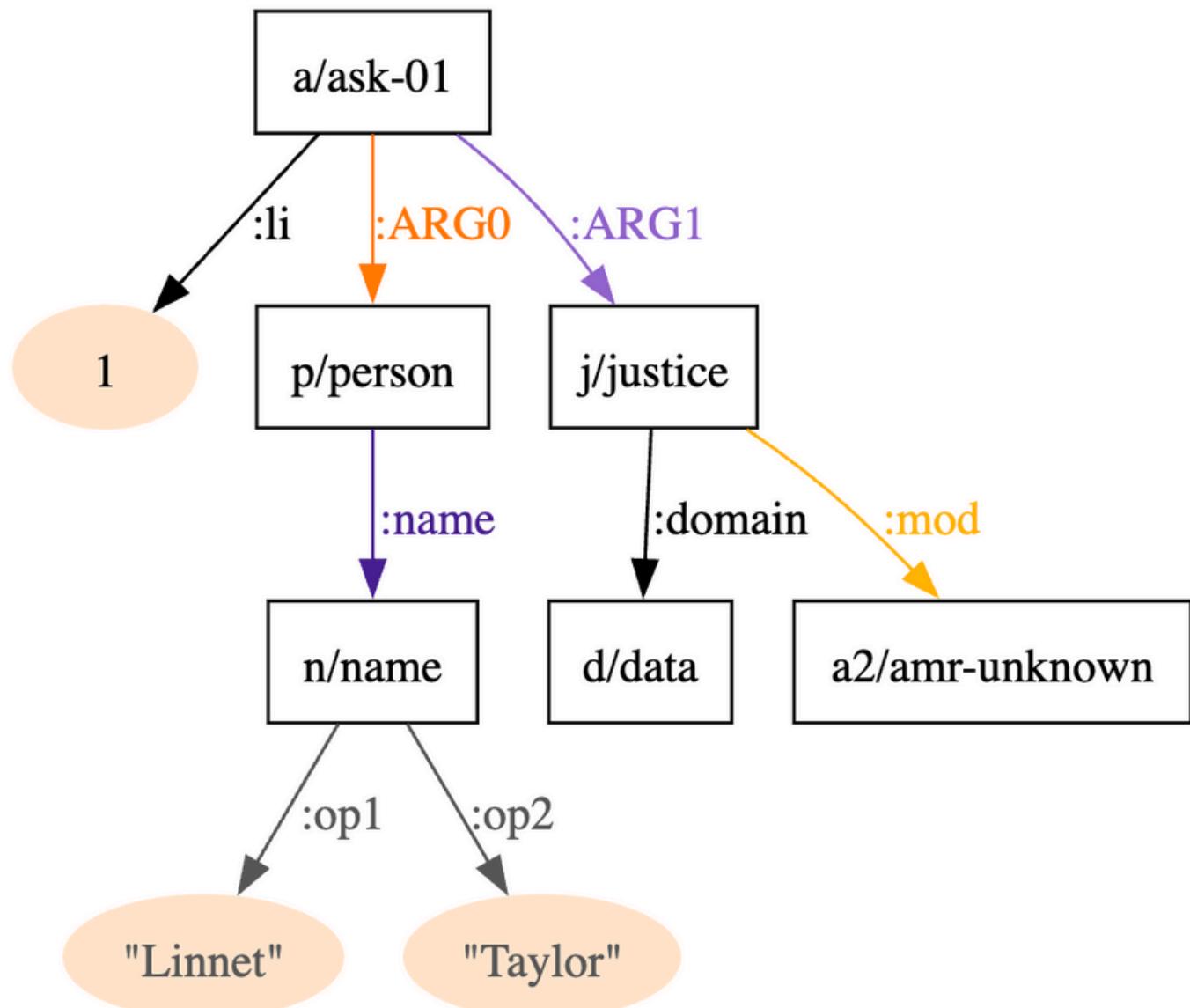
“The contrast includes court statistics and empirical research on justice institutions used by individuals in pharmacies and further studies, but this does not stop there.”

AMR graphics examples



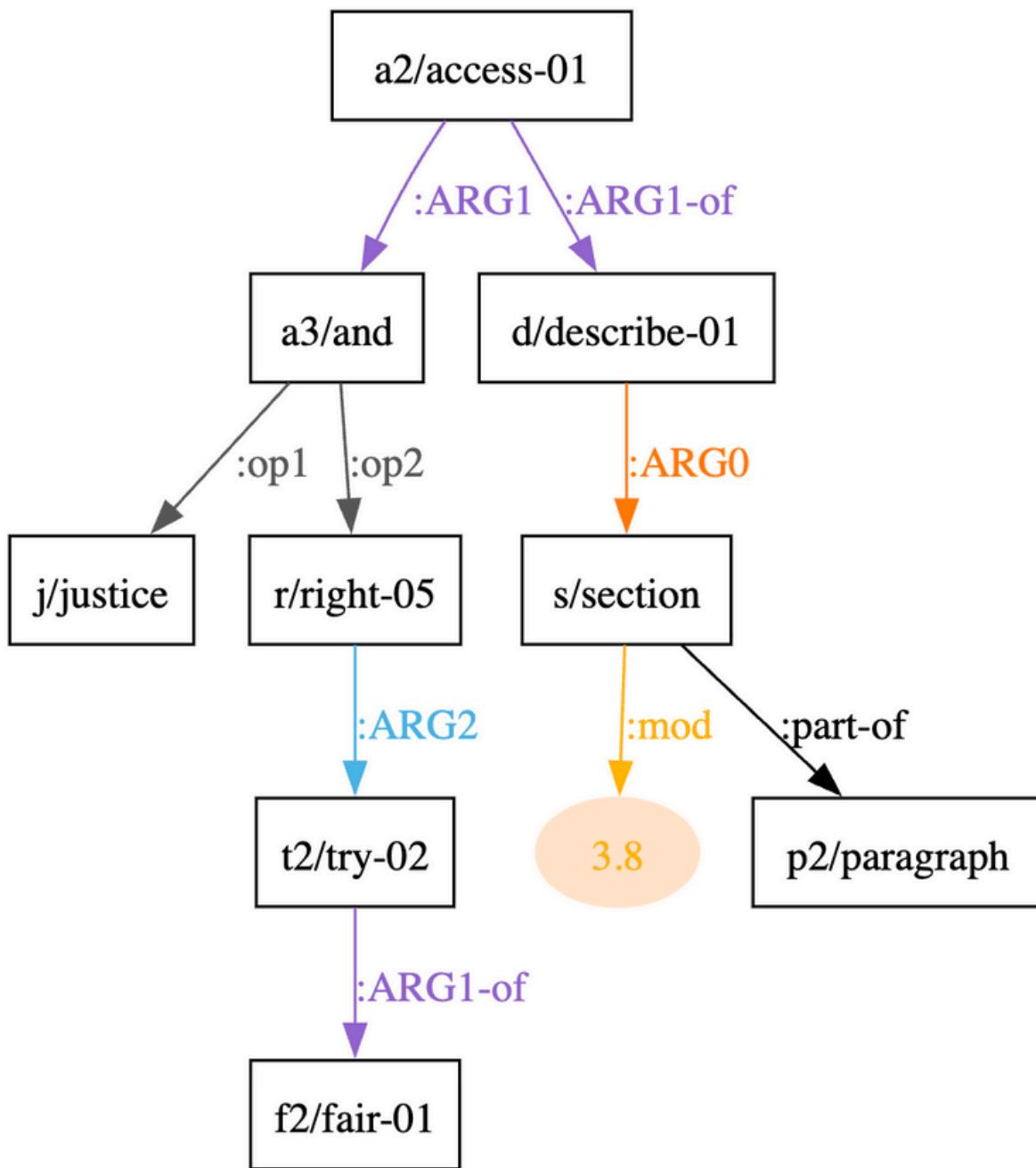
“Publication in the newspaper Indian Express: The
Dangers of a Centralised Database for Justice System.”

AMR graphics examples



"Linnet Taylor asks about justice in the field of data."

AMR graphics examples



"Section 3.8 of the paragraph describes access to justice and the right to a fair trial."

AMR Analysis of "Justice"



Different ways of using “justice”:

The way *justice* is represented in AMR depends on its role in the sentence:

- **Modifier** (`:mod`): *justice* qualifies another concept, such as a system.
Example: `(s / system :mod (j / justice))` → “justice system”
- **Conceptual argument** (`:ARG1`): *justice* is the main concept being discussed, applied, or accessed.
Example: `(a / ask-01 :ARG1 (j / justice ...))` → “asks about justice”
- **Named entity** (`:name`): *justice* is part of the name of an organization or institution.
Example: `(o2 / organization :name "Court of Justice")` → “Court of Justice”

Most of the time, when referring to an institution, justice is used as a constant string to define the name of the institution. Can we generalize rules like this one ?

AMR Analysis of "Justice"

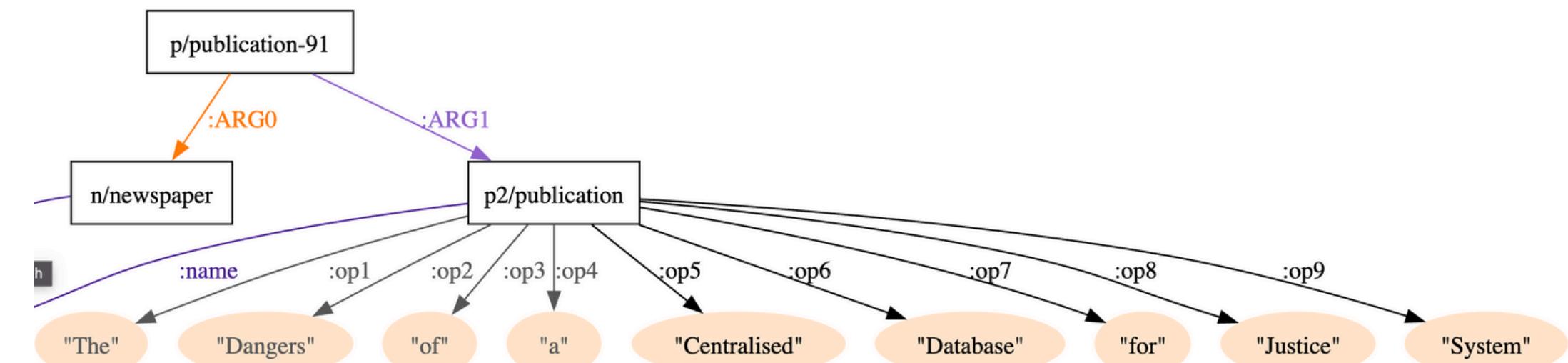
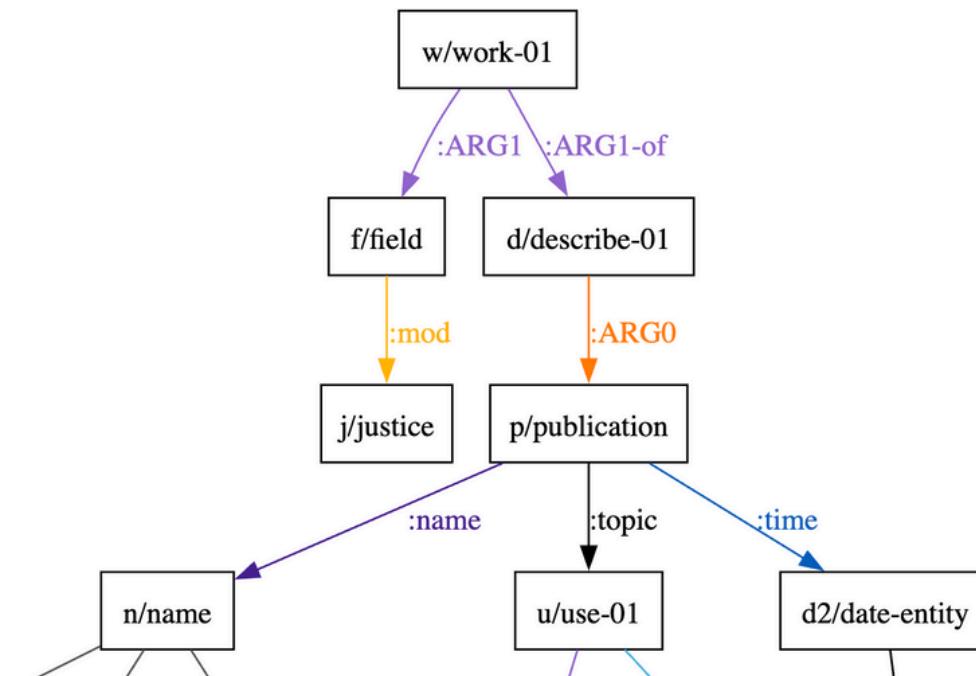
Semantic role and AMR parsing

Even when **justice** refers to the **same underlying concept**, AMR may encode it differently depending on context.

→ This shows that AMR captures contextual differences in syntax or structure, while the semantic meaning remains identical.

"The work in the field of justice is described by the publication European Ethical Charter (2018), on the use of artificial intelligence..."

"Publication in the newspaper Indian Express: The Dangers of a Centralised Database for Justice System."



Thank You

We are ready to assist you for
your next NLP project !

