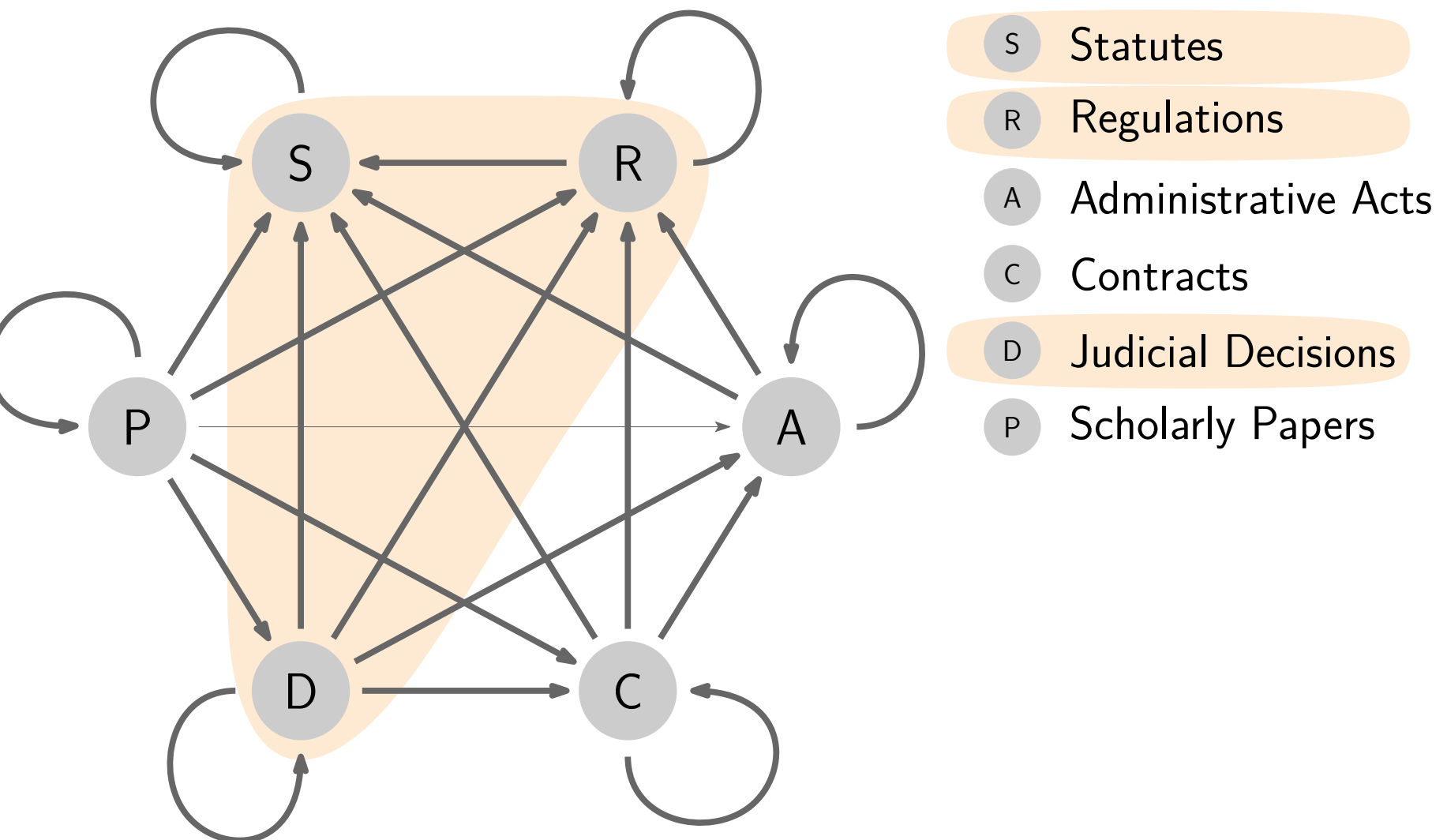# ParaGraph

Analyzing legal citation networks across document types

Titus Pünder[1] and Corinna Coupette [1,2]

1. Aalto University, Finland   2. Max Planck Institute for Informatics and Max Planck Institute for Tax Law and Public Finance, DE
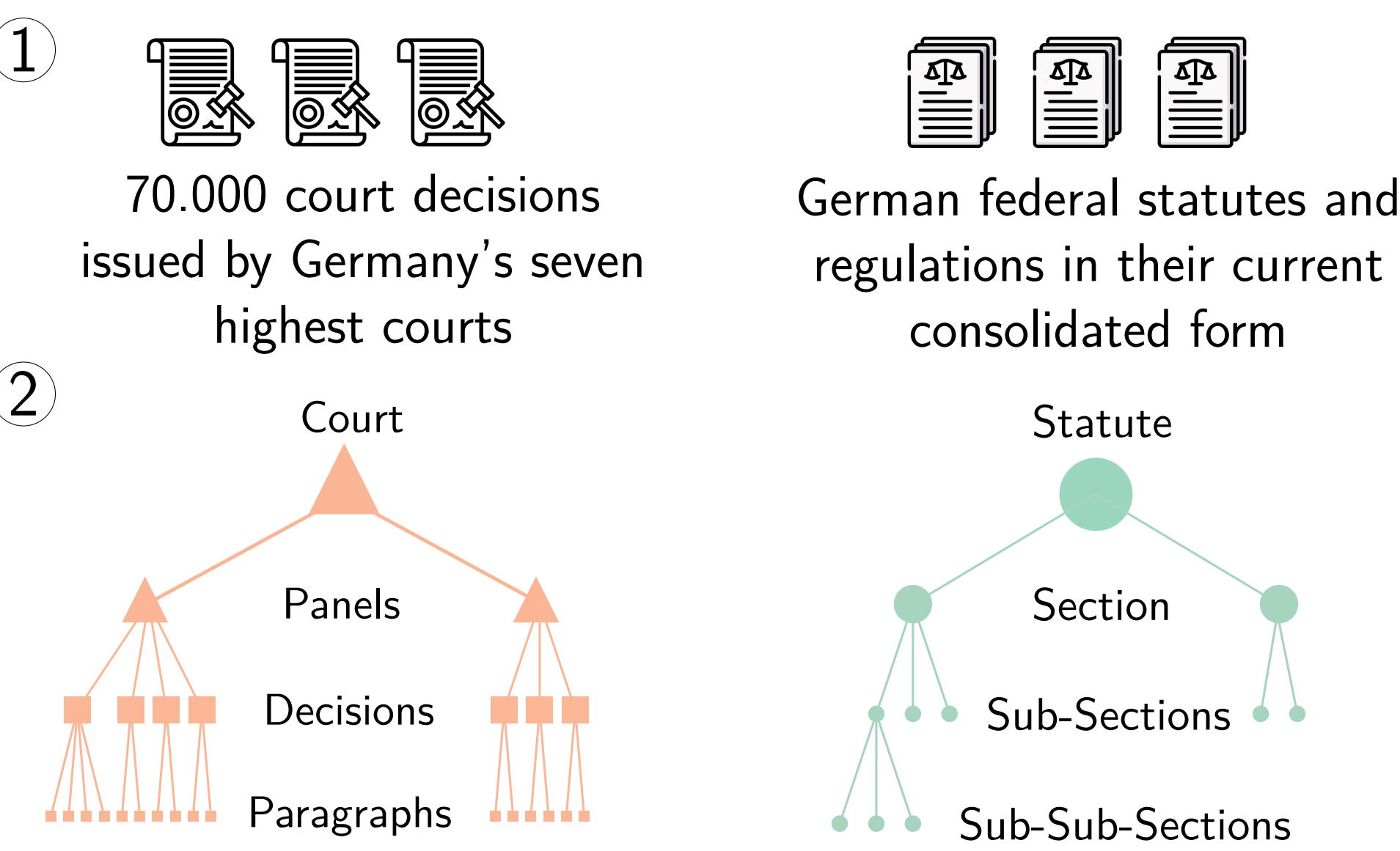
Relations between Documents

Legal systems act through people and organizations that produce *interconnected sets of documents*.

- S Statutes
- R Regulations
- A Administrative Acts
- C Contracts
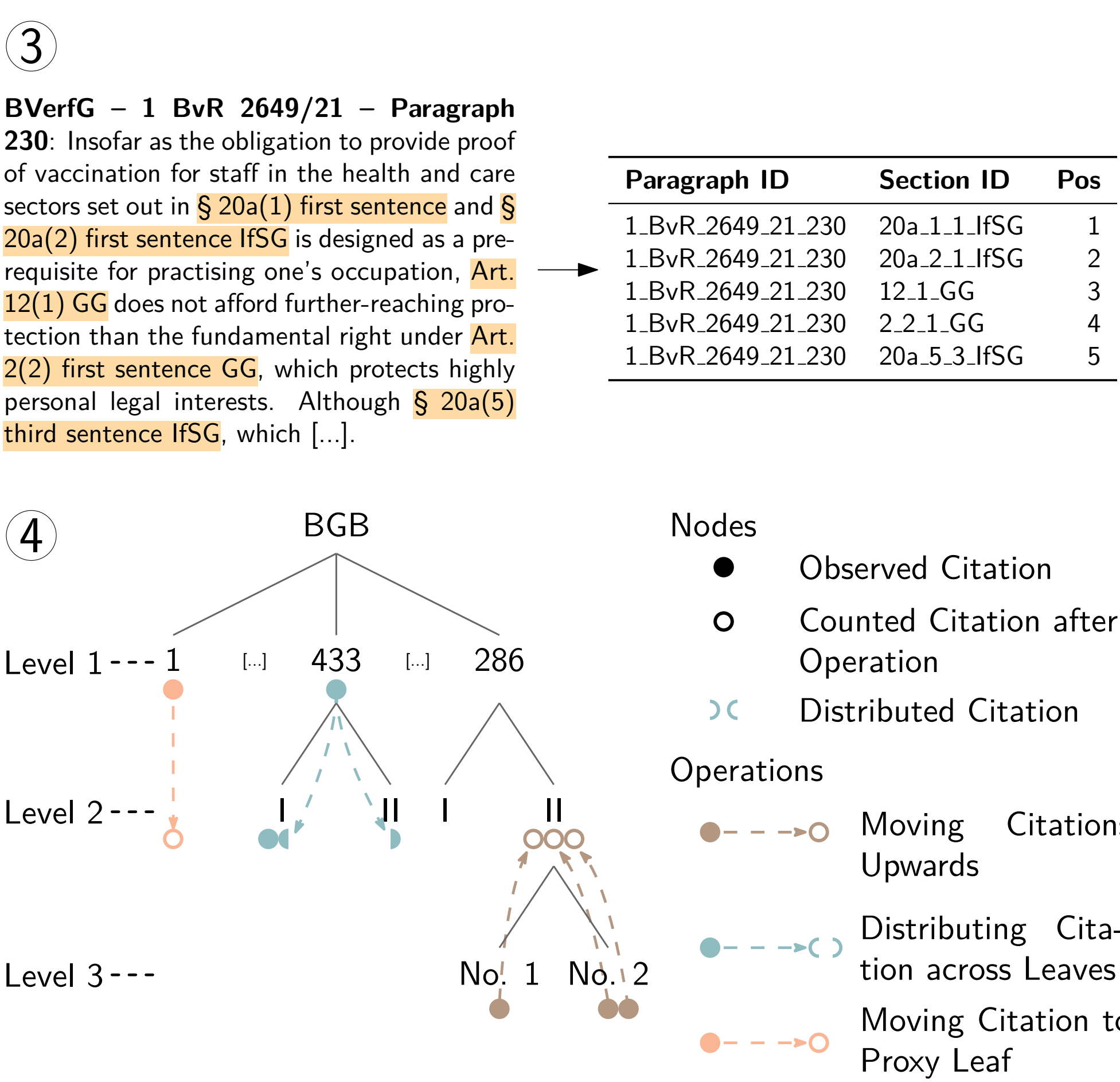- D Judicial Decisions
- P Scholarly Papers

What can we learn about statutes and decisions from studying how paragraphs of decisions reference sections of statutes?

Raw Data and Preprocessing

① 70.000 court decisions issued by Germany's seven highest courts

German federal statutes and regulations in their current consolidated form
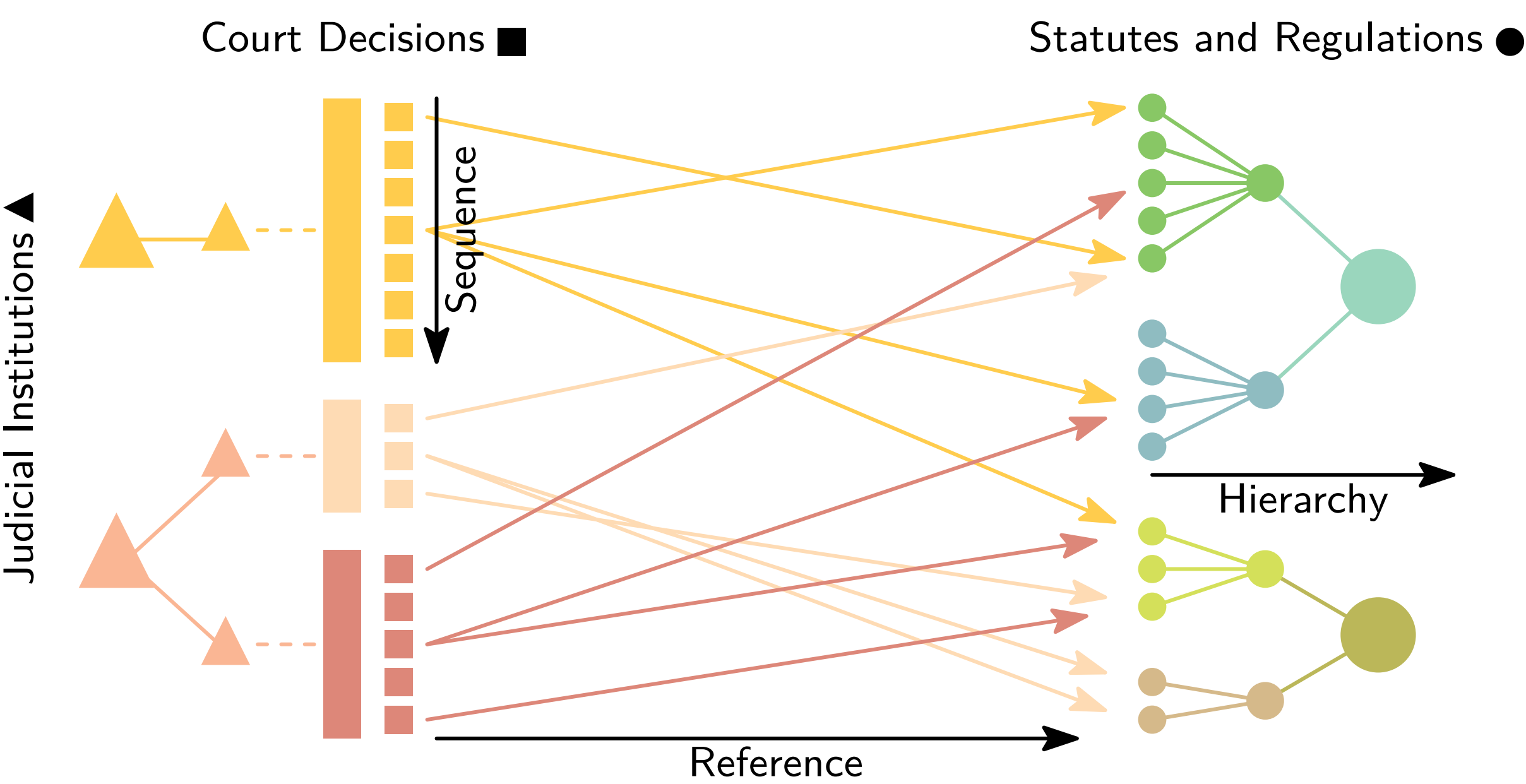
②

From the raw data ①, we extract all hierarchical metadata, including *paragraphs* and *sections* ②, as well as all atomic citations ③. Information is preserved across incoherent hierarchy levels through **dynamic counting** ④.

③ **BVerfG – 1 BvR 2649/21 – Paragraph 230**: Insofar as the obligation to provide proof of vaccination for staff in the health and care sectors set out in § 20a(1) first sentence and § 20a(2) first sentence IfSG is designed as a prerequisite for practising one's occupation, Art. 12(1) GG does not afford further-reaching protection than the fundamental right under Art. 2(2) first sentence GG, which protects highly personal legal interests. Although § 20a(5) third sentence IfSG, which [...].

Paragraph ID	Section ID	Pos
1.BvR_2649_21.230	20a_1_1.IfSG	1
1.BvR_2649_21.230	20a_2_1.IfSG	2
1.BvR_2649_21.230	12_1.GG	3
1.BvR_2649_21.230	2_2_1.GG	4
1.BvR_2649_21.230	20a_5_3.IfSG	5

④

Nodes:
- ● Observed Citation
- ○ Counted Citation after Operation
-)(Distributed Citation

Operations:
- Moving Citations Upwards
- Distributing Citation across Leaves
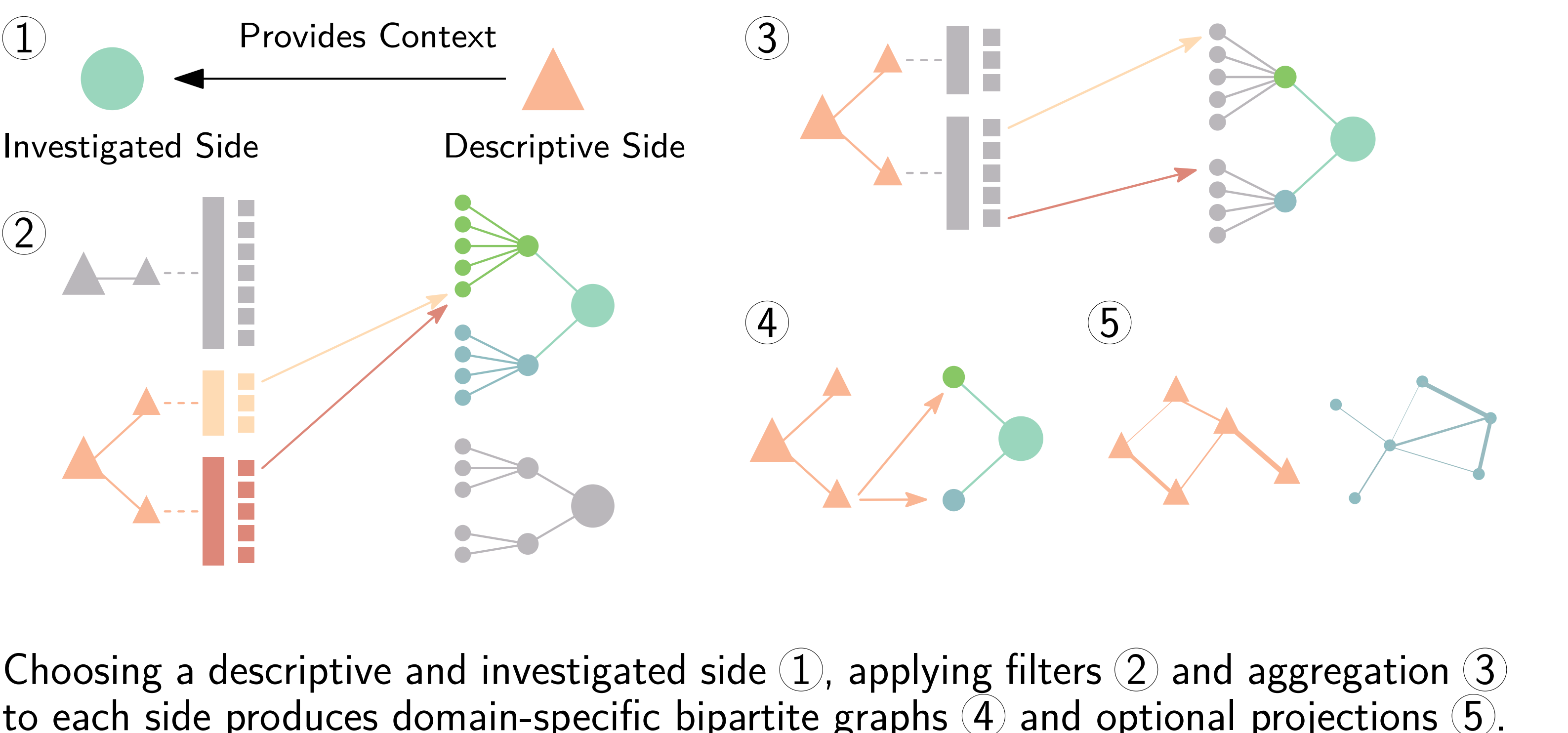- Moving Citation to Proxy Leaf

Network Representation

The reference network integrates structural and organizational metadata to provide a flexible framework that can adapt to domain-driven questions.

Analytical Approach

① Provides Context — Investigated Side ← Descriptive Side

Choosing a descriptive and investigated side ①, applying filters ② and aggregation ③ to each side produces domain-specific bipartite graphs ④ and optional projections ⑤.

Limits of Standard Metrics

When used naïvely, standard metrics such as degree centralities **reproduce data biases** and **fail to detect** substantively important sections.
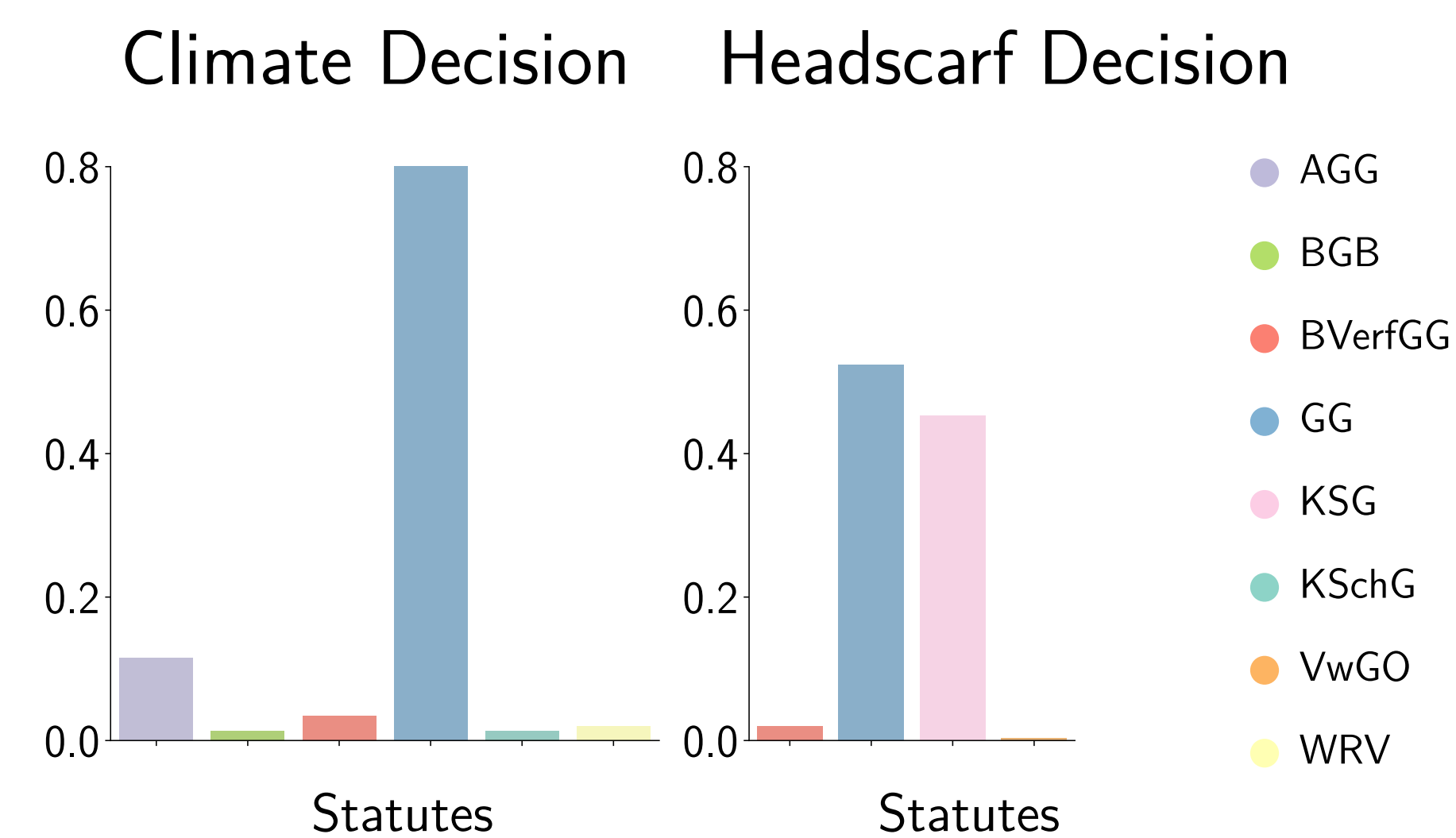
Pos	Section	Degree Centrality
1	§ 8 MarkenG	22,233
2	§ 132 VwGO	14,606
3	Art 3 GG	14,010
4	Art 103 GG	11,402
5	§ 115 FGO	7,941
6	§ 563 ZPO	7,460
7	§ 349 StPO	7,072
8	§ 126 FGO	6,865
9	Art 2 GG	6,819
10	§ 160 SGG	6,664

Edge Sets as Distributions

Incoming references characterize statute sections.

§ 242 Civil Code | Art 5 Basic Law

- BAG
- BFH
- BGH
- BPatG
- BSG
- BVerfG
- BVerwG
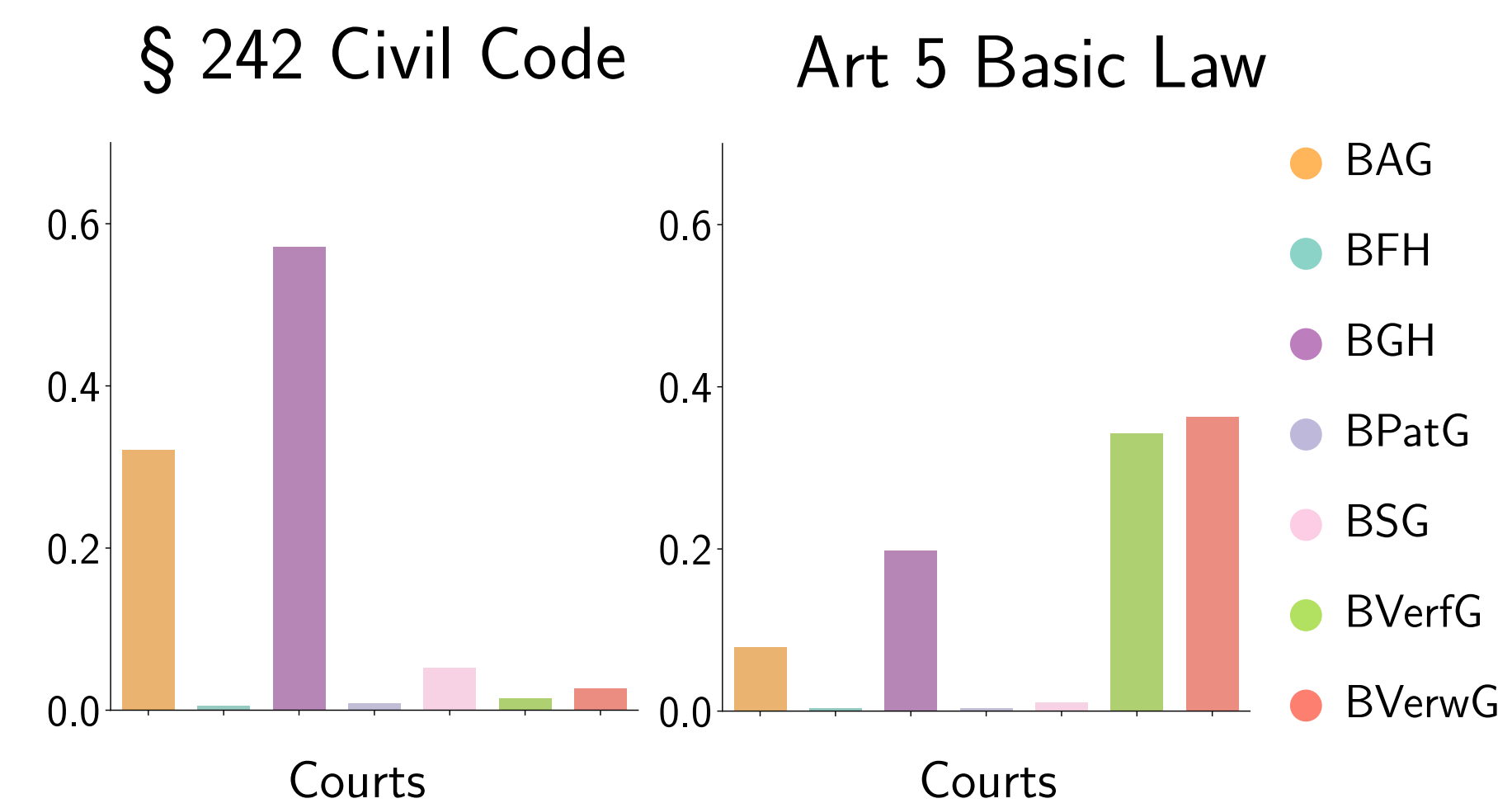
Outgoing references characterize judicial decisions.

Climate Decision | Headscarf Decision

- AGG
- BGB
- BVerfGG
- GG
- KSG
- KSchG
- VwGO
- WRV

Context-Sensitive Analysis

Co-occurrence subgraphs of sections:

①

Section Level | Section Level | Sub-Section Level

§ 823 BGB | § 823 (1) BGB | § 823 (2) BGB

- EG-FGV 2011
- BGB
- StGB
- GG
- KredWG
- ZPO
- StVG

②

§ 543 BGB | § 433 BGB | § 823 BGB

Decision Resolution

Paragraph Resolution

- EG-FGV 2011
- BGB
- StGB
- GG
- KredWG
- ZPO
- StVG

③

Constitutional Court | Civil Court | Penal Court

Art 5 Basic Law | Art 5 Basic Law | Art 5 Basic Law
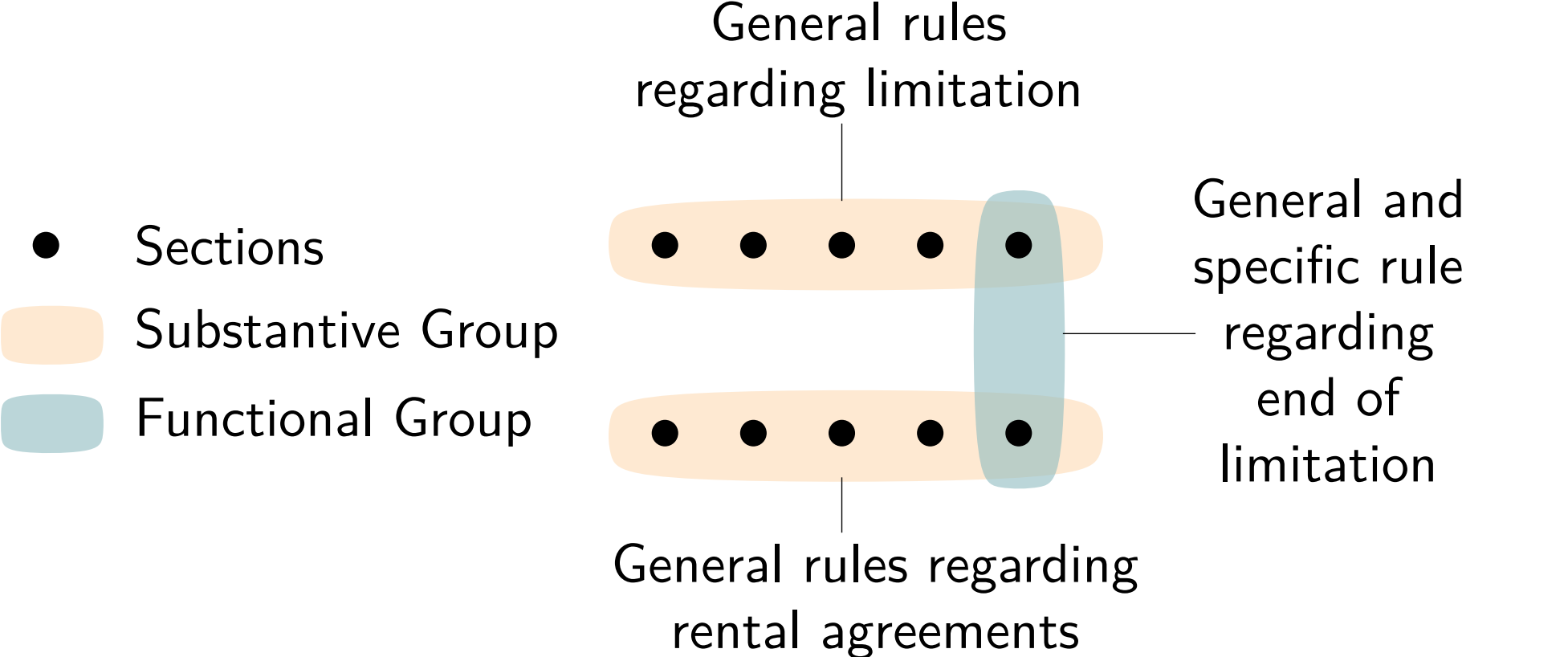
- BVerfGG
- GG
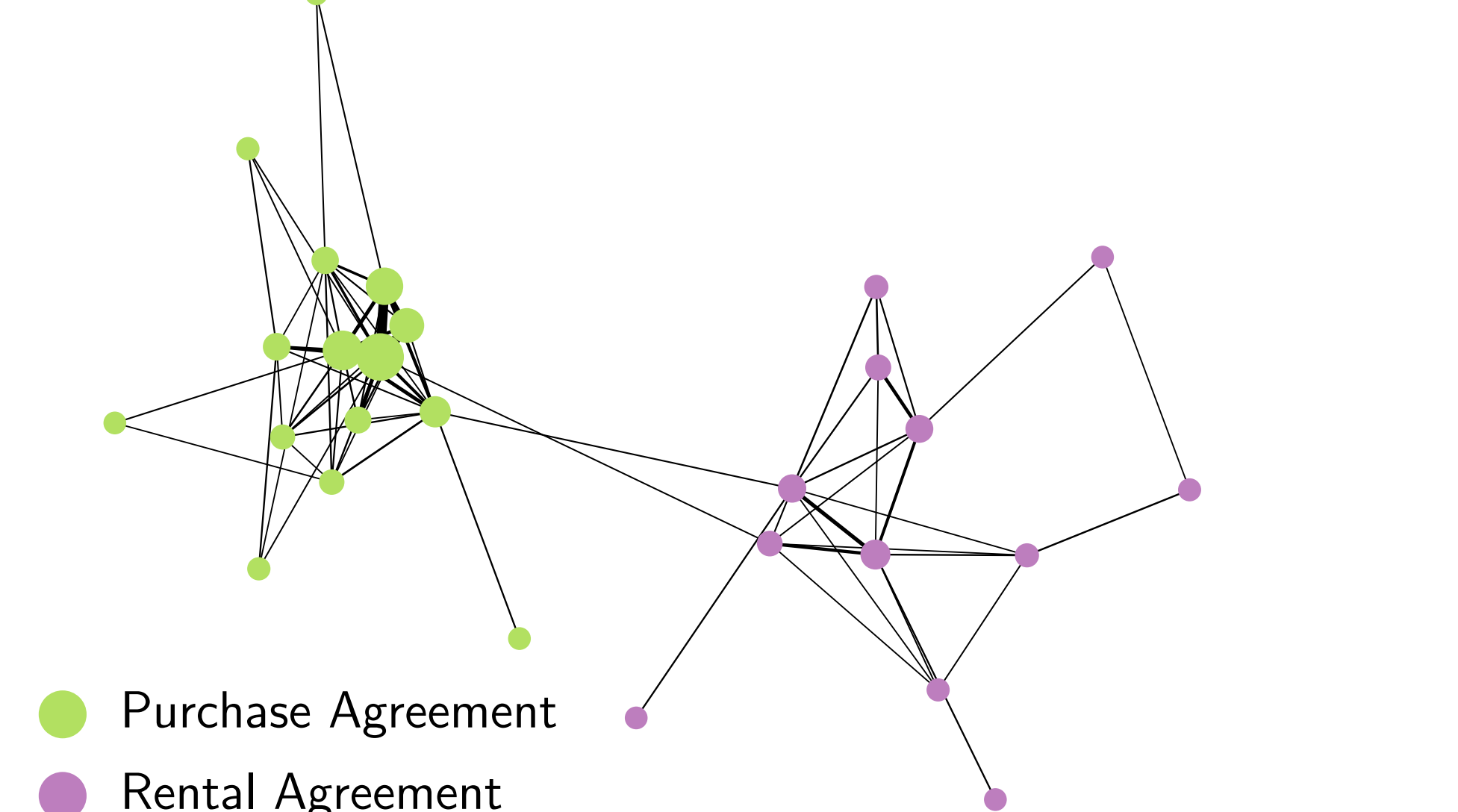- StGB
- StPO

Nodes Sizes and Edge Widths Min-Max-Scaled
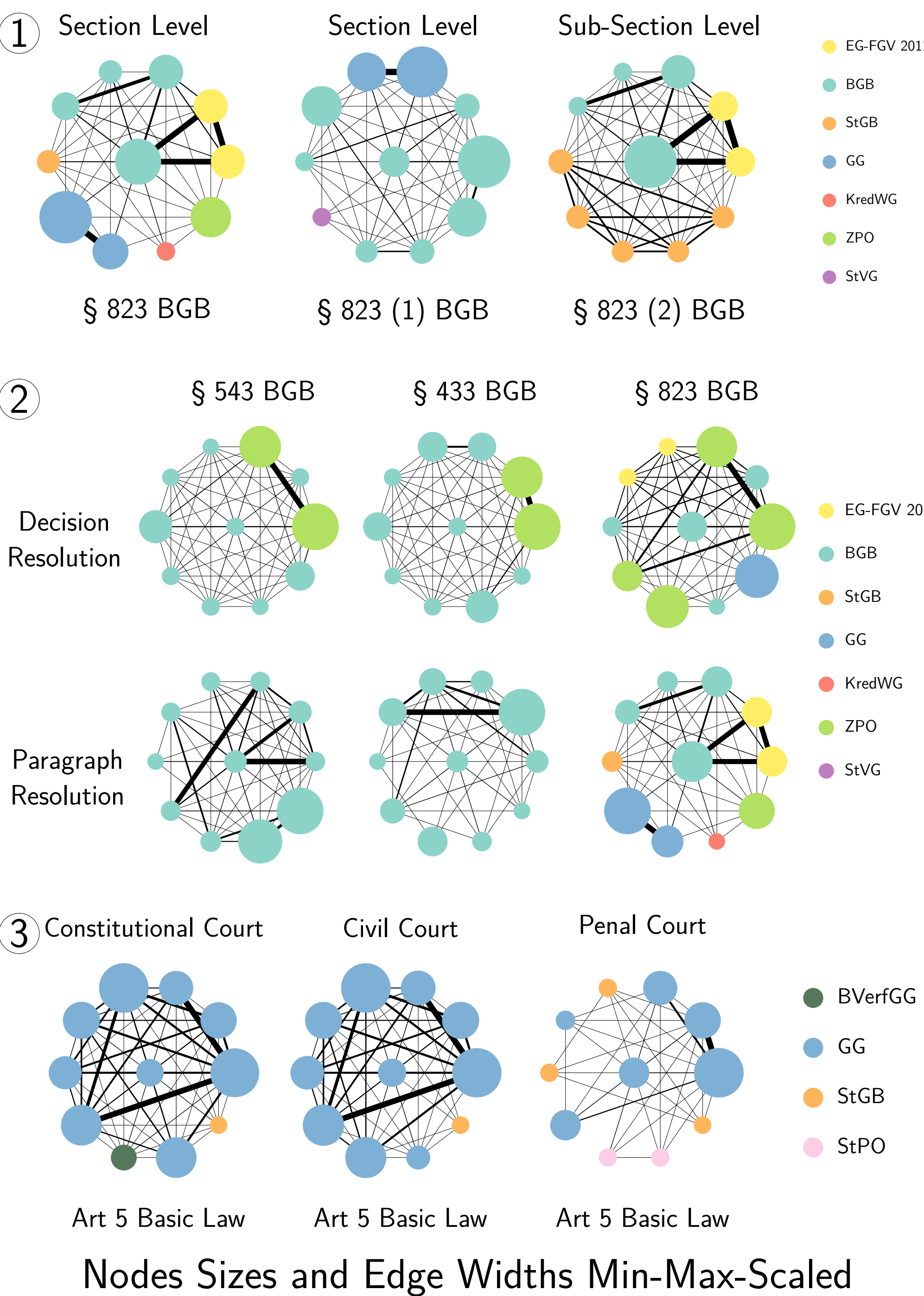
Controling the resolution level of the descriptive side ① or the investigated side ② increases differential power and reduces noise. Filtering the underlying data ③ characterizes sections within different legal contexts.

Groups of Sections

Legal categorization of section groups:

General rules regarding limitation

General and specific rule regarding end of limitation

General rules regarding rental agreements

- ● Sections
- Substantive Group
- Functional Group

Network representation of two substantive groups:

- Purchase Agreement
- Rental Agreement

Limitations and Next Steps

- Expand data set vertically and horizontally to tackle *incomplete and biased datasets*.
- Include inter-statute references and inter-decision citations.
- Leverage NLP methods to extract implicit references.
- Model the dataset as a hypergraph to address *shortcomings of pairwise projections to capture reference groups*.