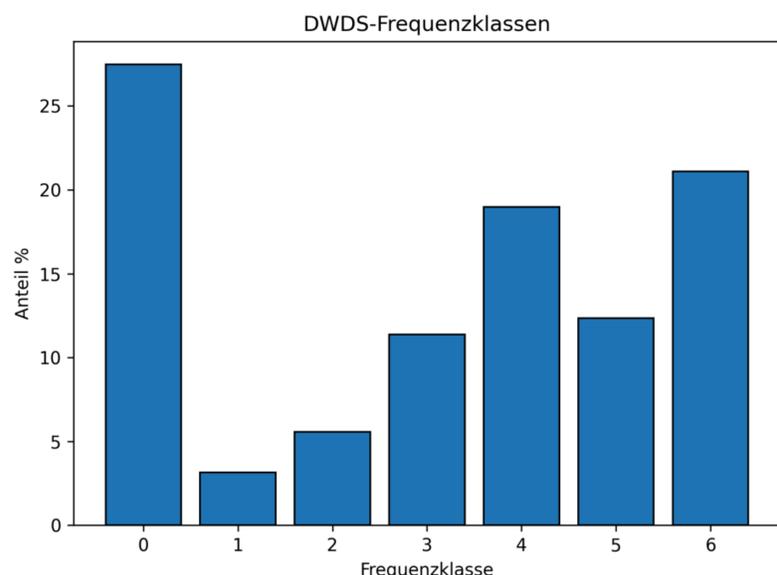


## Abstract

Die Arbeit untersucht die **lexikalische Komplexität deutscher Verwaltungstexte** anhand von über 3.000 Leistungsbeschreibungen. Mithilfe von NLP und statistischen Metriken wurden **Wortvielfalt**, **Nominalisierungen** und **Lesbarkeit** analysiert.

**Ergebnis:** Verwaltungssprache ist stark verdichtet, nutzt viele seltene Wörter und erschwert das Verständnis. Vereinfachte Sprache kann die Zugänglichkeit amtlicher Informationen deutlich verbessern.

## Lexikalische Sophistication



**DWDS-Frequenzklassen:** 0 = sehr selten, 6 = sehr häufig.

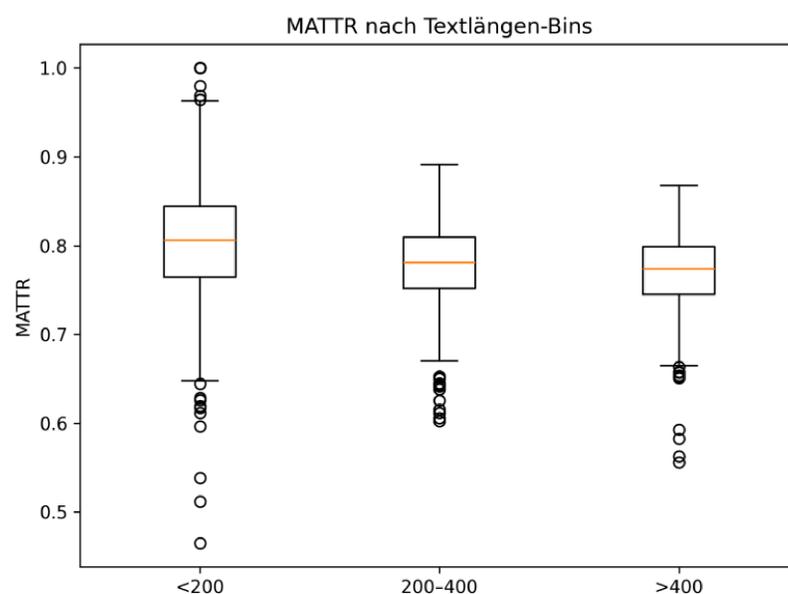
Die **Verteilung** weist einen substantiellen Anteil seltener Lexik auf, was auf die Verwendung **registerspezifischer Terminologie** und **nominaler Derivate** hindeutet.

Einordnung der **Lesbarkeitsresultate:** Ein höherer Anteil seltener Lexik korrespondiert mit erhöhten Anforderungen an **Wortschatztiefe** und **Bedeutungserschließung**.

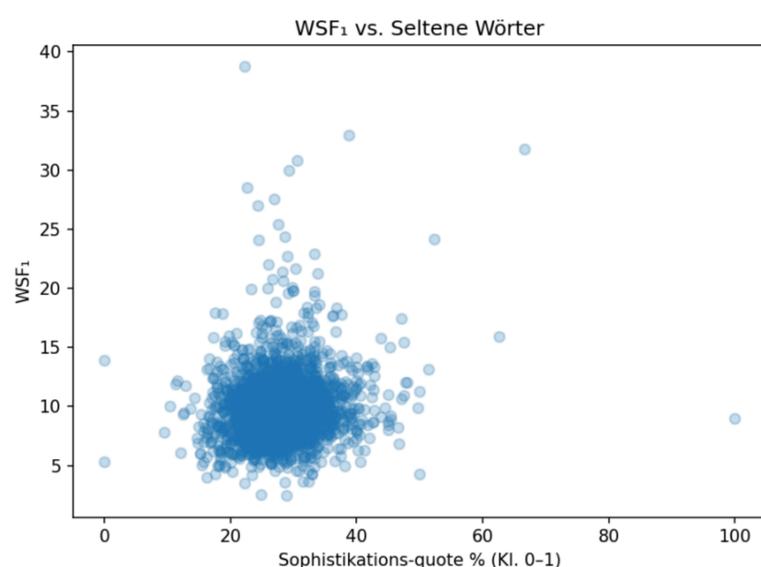
## Lexikalische Vielfalt

**MATTR:** Moving Average Type-Token Ratio. Eine quantitative Messung der **lexikalischen Vielfalt** in einem Text.

Ein **hoher MATTR-Wert** deutet auf eine **große Vielfalt** an unterschiedlichen Wörtern (Typen) im Verhältnis zur Gesamtzahl der Wörter (Tokens) hin.



## Lesbarkeit



Der Streudiagrammvergleich zwischen **WSF1** (Lesbarkeitsindex für Sachtexte) und den Anteilen seltener **DWDS-Wortklassen** zeigt eine systematische Kopplung: Mit zunehmender **Seltenheit** im Wortschatz steigen die **Anforderungen** an **Dekodierung** und Satzverstehen.

Der Zusammenhang bei Texten mit **hoher Nominalisierungsrate** zeigt, dass seltene Lexik ein zentraler Treiber **sinkender Lesbarkeit** in Verwaltungssprache ist.