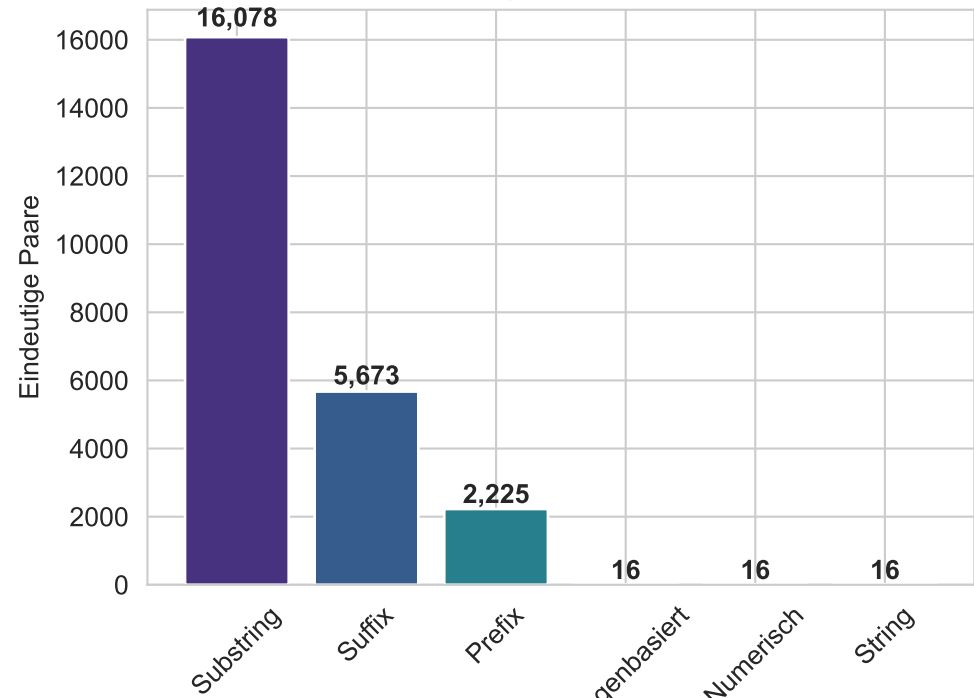


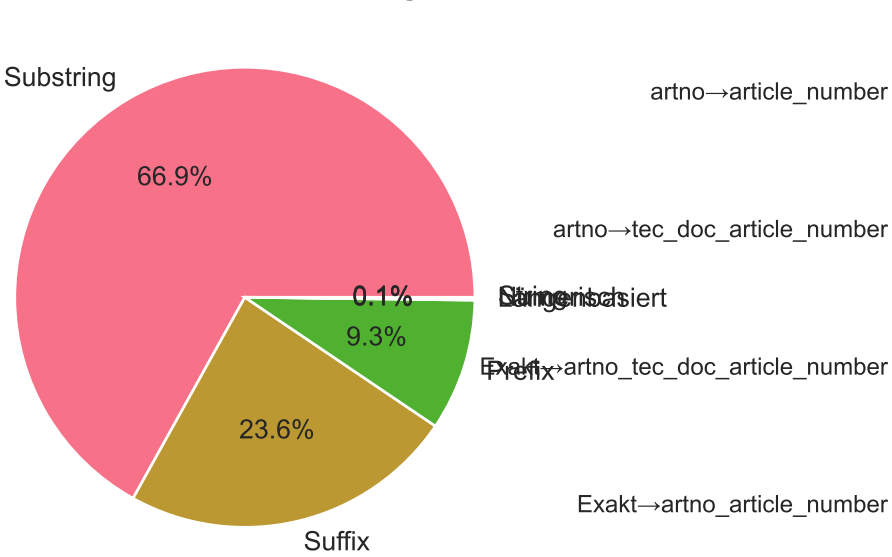
DETERMINISTIC MATCHING DASHBOARD (DUPLIKATFREI)
TecDoc vs CMD TMD - Wissenschaftliche Analyse ohne Duplikate

- GESAMT-STATISTIKEN:
- 24,024 eindeutige (TecDoc ↔ CMD) Paare gefunden
 - 3,936 eindeutige TecDoc-Werte beteiligt
 - 21,180 eindeutige CMD-Werte beteiligt
 - Beste Methode: Substring | Beste Spalten: artno → article_number
 - Sample: 50k Zeilen von 40GB (0.1% repräsentativ)

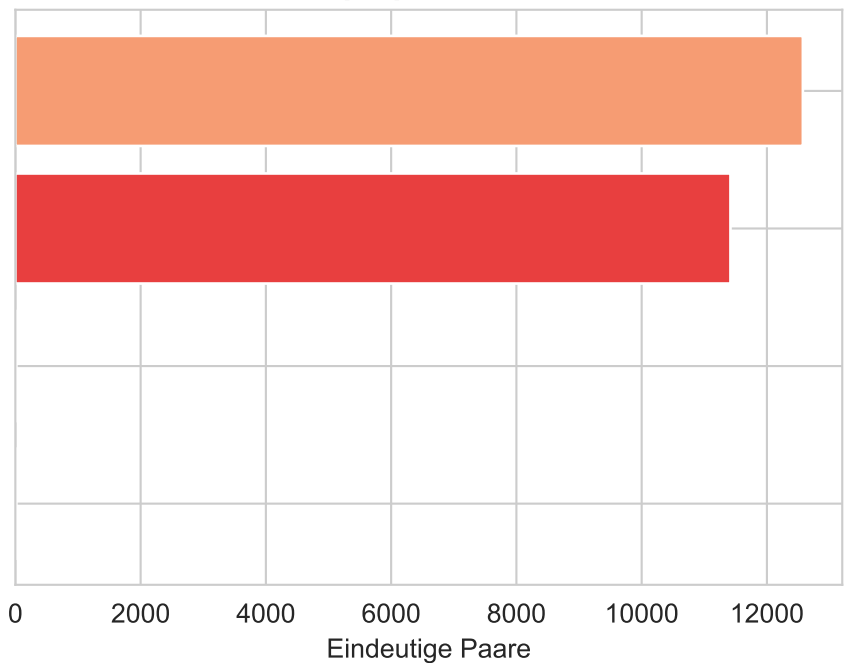
Top Methoden



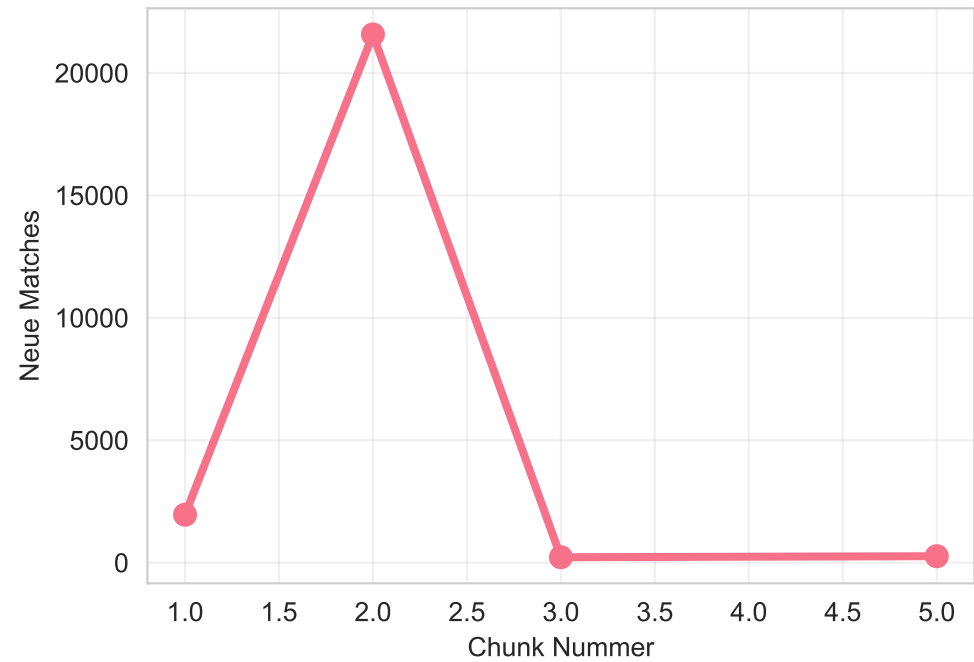
Methoden-Verteilung



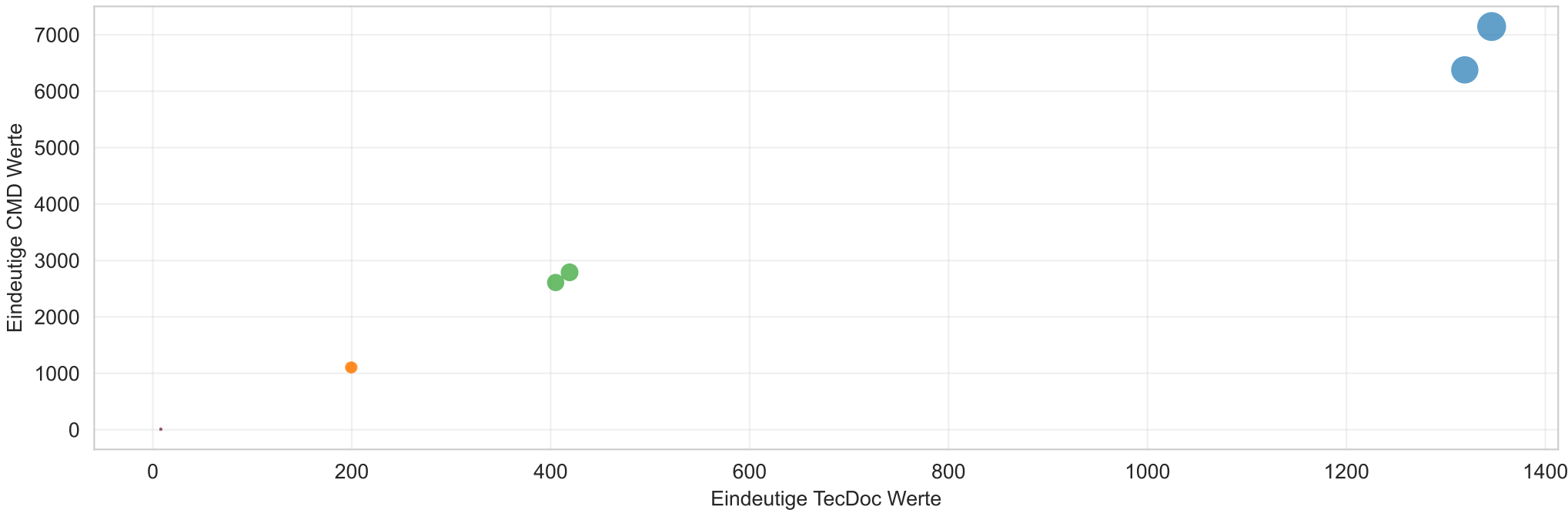
Top Spalten-Paare



Fortschritt pro Chunk



Methoden-Effizienz (Bubble-Size = Paare)



- Substring
- Prefix
- Suffix
- Numerisch
- Längenbasiert
- String