



Automatisiertes Named Entity Linking – Ein Feldversuch

Von API-Abfragen zur
hybriden Vektorsuche

Stephan Makowski · 21.10.2026



BERGISCHE
UNIVERSITÄT
WUPPERTAL

Ein Name führt zu mehreren Kandidaten

Eine Bezeichnung

**Registereintrag
„Grefrath“**

Unterschiedliche Orte

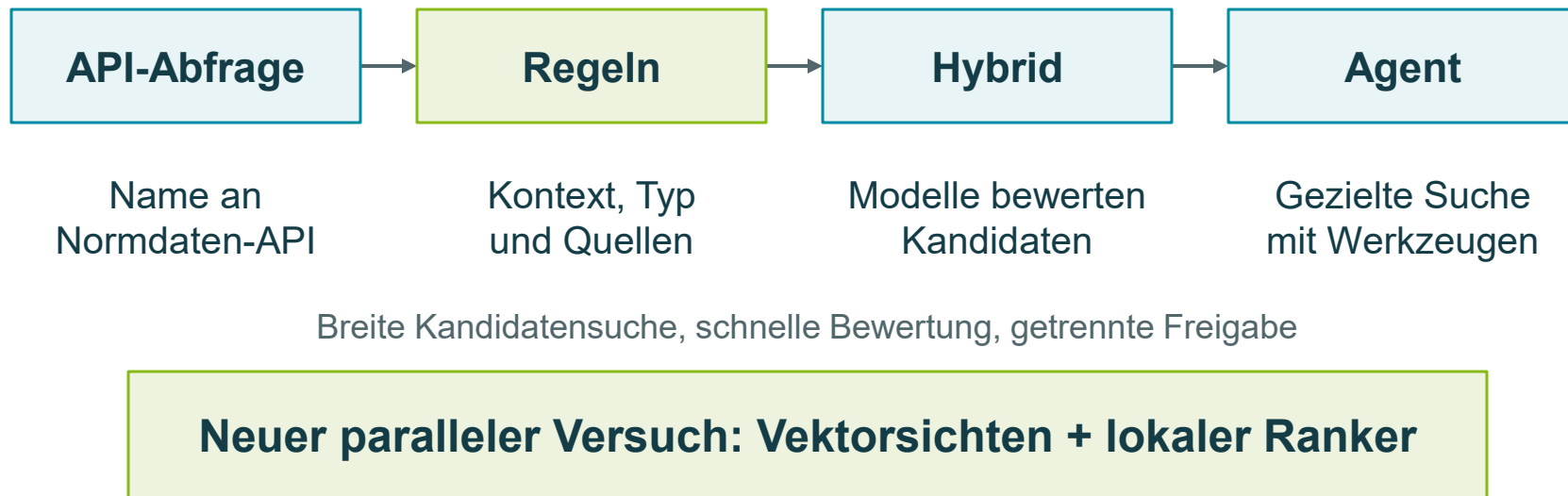
Grefrath bei Neuss

Grefrath im Kreis Viersen

Grefrath bei Frechen

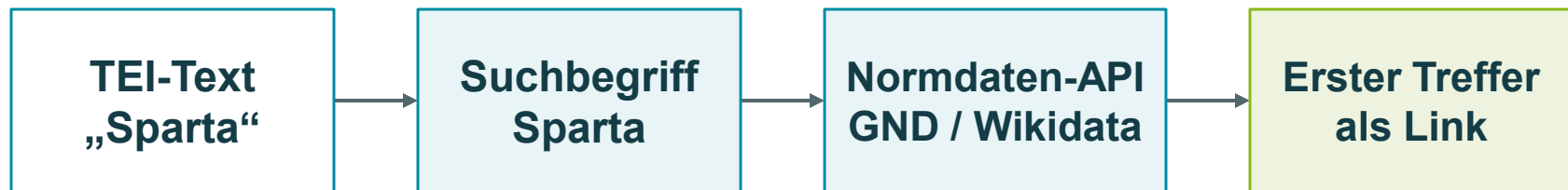
Namen, Notizen und Textbelege helfen bei der Unterscheidung. Eine passende Schreibweise allein reicht nicht.

Wie sich der Ansatz entwickelt hat



Die Entwicklungsstufen sind Funktionsmodelle. Ein größerer Agent ersetzt die lokale Verarbeitung nicht automatisch.

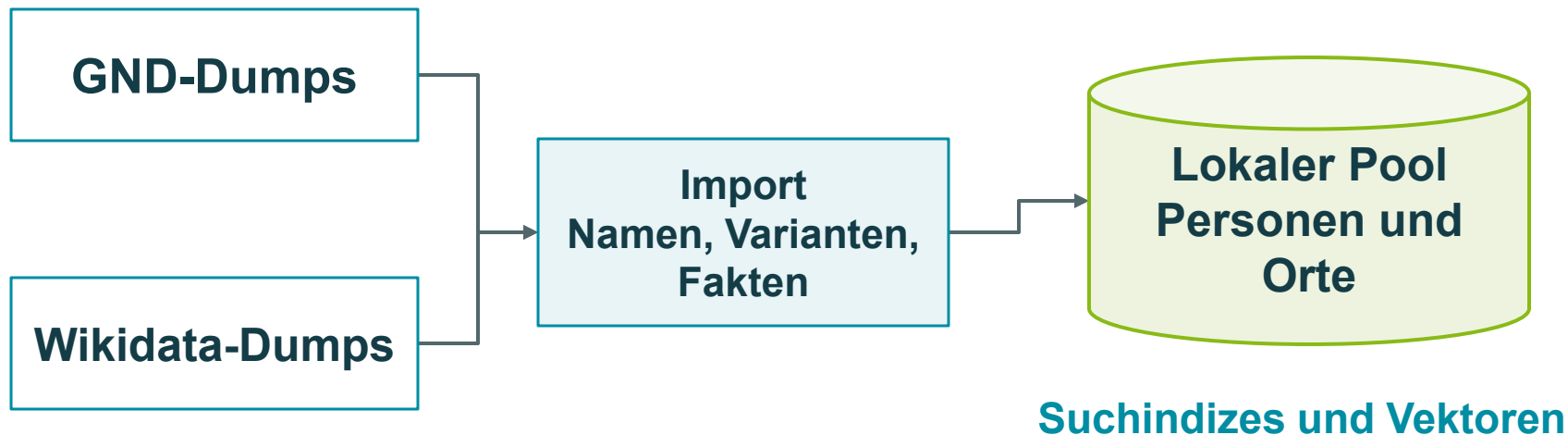
Stufe 1: Eine einfache Normdatenabfrage



Risiko: Der erste API-Treffer kann einen anderen Ort bezeichnen.

Eine API liefert Suchtreffer. Ihre Reihenfolge bestätigt noch keine Identität.

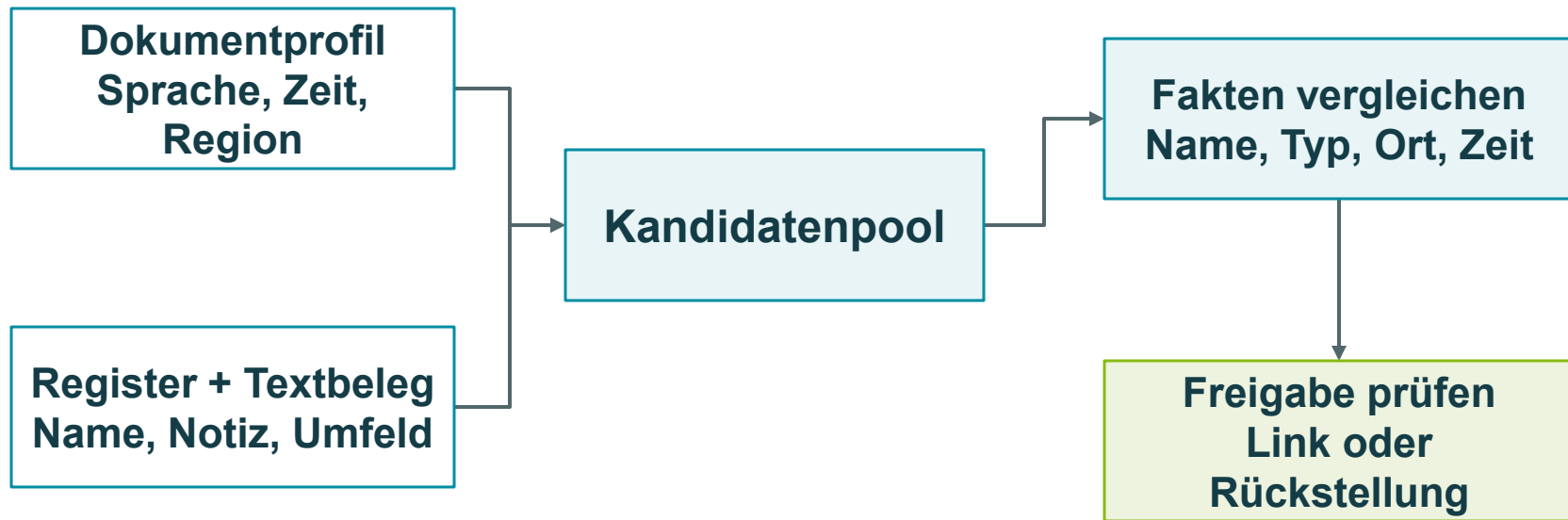
Unser eigener Pool schafft eine stabile Suchbasis



Wiederholbare Suche ohne Einzelabfrage jeder Normdaten-API

API-Ausfälle und Drosselung blockieren die lokale Suche nicht. Aktualisierung und Importqualität bleiben notwendig.

Stufe 2: Kontext und Regeln grenzen die Suche ein



Regeln machen Entscheidungen nachvollziehbar. Historische Zusätze und fehlende Angaben brauchen eine eigene Behandlung.

BERT und LLM übernehmen unterschiedliche Aufgaben

BERT / lokaler Ranker

**Register und Kandidat
als Eingabepaar**



**Passung bewerten
Kandidaten neu ordnen**

LLM, beispielsweise Qwen

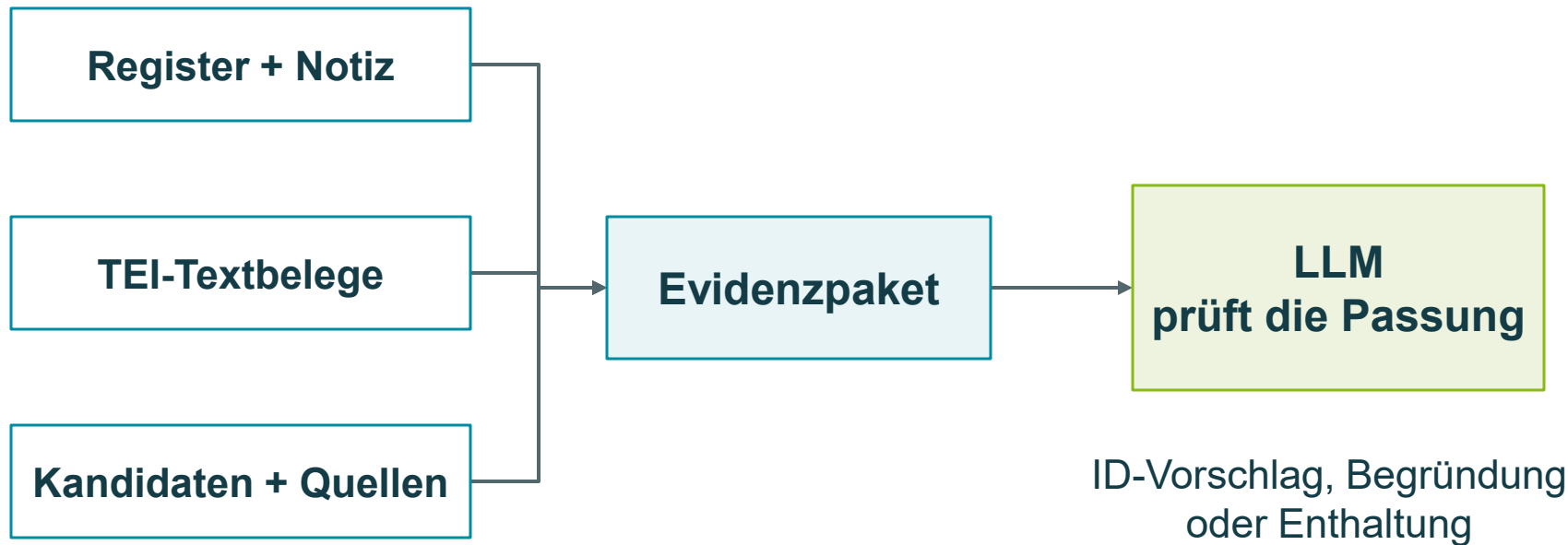
**Register, Textbelege
und Kandidatenliste**



**Widersprüche erklären
ID vorschlagen oder enthalten**

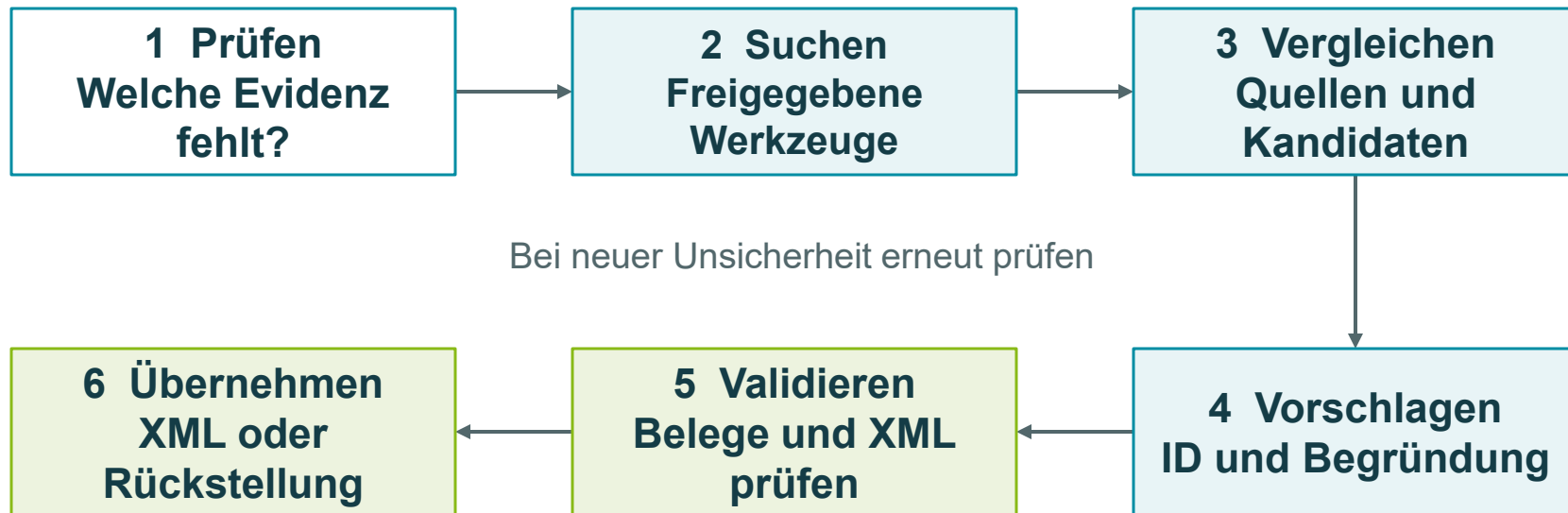
BERT recherchiert nicht selbst. Ein LLM benötigt freigegebene Werkzeuge für zusätzliche Suche.

Stufe 3: Das LLM prüft ein Evidenzpaket



Die Pipeline prüft die Ausgabe vor der XML-Übernahme. Referenz-IDs bleiben in Tests ausschließlich in der Auswertung.

Stufe 4: Ein Agent kann gezielt nachrecherchieren



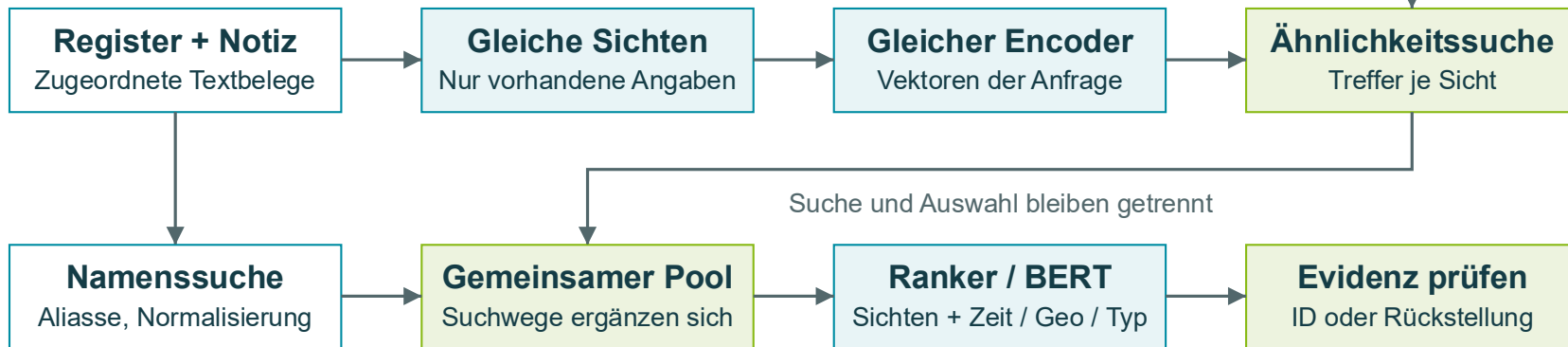
Optionaler Rechercheпад. Goldstandard-Vergleiche gehören in die Offline-Auswertung, nicht in das Agentenwissen.

Der neue Versuch: mehrteilige Vektorsuche

Einmalig: Normdaten vorbereiten



Je Anfrage: Register und Text mit dem Index vergleichen



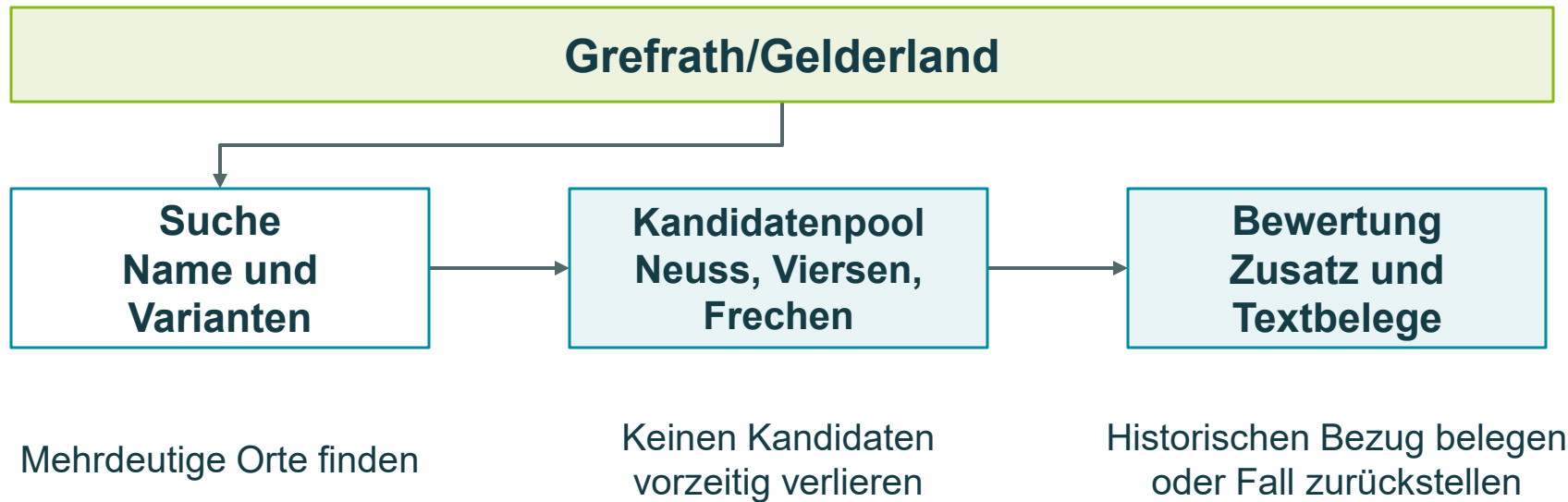
Ähnlichkeit ist keine Identitätswahrscheinlichkeit. Lebensdaten und Koordinaten bleiben zusätzlich strukturierte Fakten.

Jede Sicht beantwortet eine andere Frage

Sicht	Personen	Orte
Namen	Name und Varianten	Name und historische Namen
Profil	Beruf, Ämter, Biografie	Ortstyp und Beschreibung
Umfeld	Institutionen, Beziehungen	Region und Nachbarorte
Zeit und Raum	Lebensdaten, Wirkungsort	Geltungszeit und Koordinaten

Text erhält Vektoren. Lebensdaten und Entfernungen bleiben zusätzlich exakt prüfbare Fakten. Fehlend heißt nicht falsch.

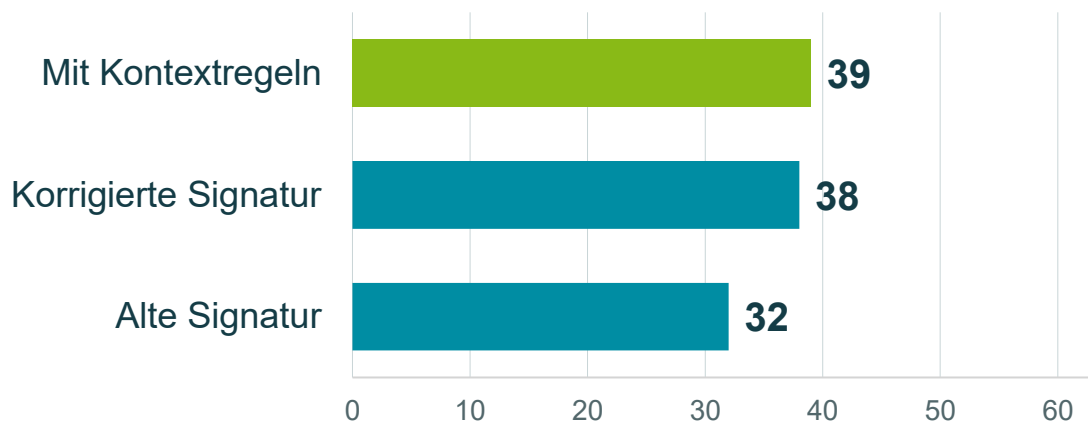
Grefrath: Suche und Bewertung ergänzen sich



„Gelderland“ bleibt als Zusatz erhalten. Vektorähnlichkeit allein beweist die historische Zuordnung zu keinem Kandidaten.

Unabhängiger Ortstest: Suche vor Auswahl

DER STURM: 81 Orte und 174 Briefe



61 / 63

Referenzziel im Pool
96,83 %

39 / 63

strikt auf Rang 1

63 lokal vorhandene, streng verknüpfte Referenzen. Ortsbezogen: 49/63 nach Einzelfallprüfung, andere Bewertungsdefinition.

Fallstricke bleiben auch bei guten Vektoren

Registerangabe

Name oder Quellen-ID widersprechen dem Text.

Normdatensatz

Lebensdaten oder Verwaltungsfelder sind fehlerhaft.

Modellierung

Stadt, Siedlung und Verwaltungssitz haben eigene IDs.

Historischer Bezug

Heutige Gebiete decken frühere Räume nicht exakt ab.

GND und Wikidata gelten bei belegter Identität gleichwertig. Ähnliche Namen oder Koordinaten allein sind kein Identitätsbeweis.

Arbeitsstand und nächste Prüfungen



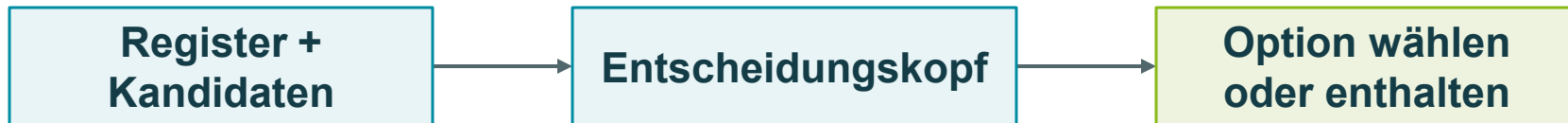
Noch nötig

Zusätze durchgängig auswerten, Kontext gezielt nutzen und Gewinne gegen Regressionen prüfen.

Modellversuche: BERT, Jina, Kev und Clef. Jev bislang recherchiert. LLM-Recherche bleibt eine optionale Ergänzung.

Jev, Kev und weitere Entscheidungsmodelle

Prinzip bei Kev und Clef: Auswahl aus vorgegebenen Optionen, ohne freien Antworttext



Variante	Referenztreffer	Aktive Zeit
Qwen 27B remote, angereichert	140 / 143	7,04 min
Kev-9B lokal, kompakt	139 / 143	26,61 min
Clef 27B lokal, kompakt	138 / 143	49,39 min

Jev: recherchiert, noch kein eigener Benchmark.

Hohe Auswahlquote, lokal hier kein Laufzeitvorteil.

150 Fälle, 143 bewertbar, gleiche Top-10-Pools. Nur Entscheidungszeit, ohne Laden und Suche.
Bekannte Testdaten, unterschiedliche Profile und Hardware. Keine produktive Ablösung belegt.