

Semantische Technologien für das BIOfid-Portal

Adrian Pachzelt

Universitätsbibliothek J. C. Senckenberg

Fachinformationsdienst
Biodiversitätsforschung



Fagus sylvatica

Suchen

 Stichwort Semantisch

Dokumente

37 Suchergebnisse

Angezeigte Trefferzahl: 10

Mitteilungen über die Eibe, besonders über die Dichtigkeit ihres Auftretens

Download als PDF

Datenbank: Dig. Samm. UB JCS Dokumentenart: Artikel Sprache: ger Publikationsjahr: 1912 Autor: Conwentz, Hugo

Zeitschrift: Bericht über die Zusammenkunft der Freien Vereinigung für Pflanzengeographie und Systematische Botanik

▼ Seitentreffer

Seite 46 (1 Treffer):

...Der Bestand setzt sich hauptsächlich aus Fichte, Tanne und Rotbuche mit Eibenunterwuchs zusammen. Kollmann aus Weilheim, des eigentlichen Entdeckers dieses Standortes....

Seite 48 (3 Treffer):

...Die Eibe kommt als Unterwuchs in dem aus Fichte und Rotbuche gebildeten Hauptbestand, sowie auch frei in den Felspartien der Donau vor. Auf einer Gesamtfläche von 60 ha stehen, nach Schätzung des Forstamts, etwa 600 Eiben), d. h. 10 Stück per ha....

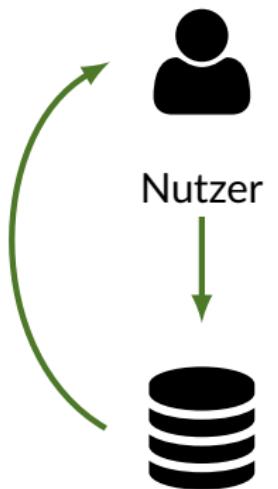
...Rotbuche, Kiefer und Eiche bilden den Hauptbestand, in welchem die Eibe mit Wacholder und Hasel als Unterholz auftritt....

...In den Jagen 139c, 140b und 141a findet sich noch der ursprüngliche Bestand, in welchem die Rotbuche vorherrscht, stellenweise auch rein vorkommt....

Gefundene Taxa

Taxon (allgemein)	Tier	Pflanze	Ort	Zeit
-------------------	------	---------	-----	------

Was bisher geschah...



Solr

elasticsearch

Dokumentendatenbank

Probleme der reinen Indexsuche

- Gesuchtes Wort muss im Dokument vorkommen
- Kuration von Synonym-Listen
- in biologischen Texten: *Fagus* \notin *F. sylvatica*

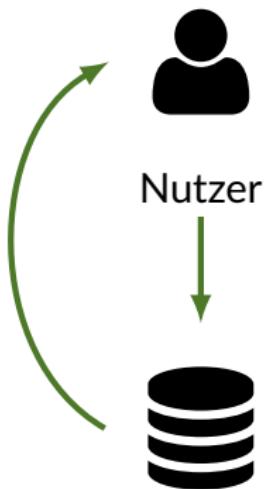
Wie können diese Probleme
behoben werden?

Definition

Durch die Verwendung von **Hintergrundwissen** wird bei einer semantischen Suchmaschine die **inhaltliche Bedeutung von Texten und Suchanfragen** berücksichtigt. Es wird nicht nur nach Wörtern im Text, wie bei Keyword-basierten Suchmaschinen, gesucht. [...]

Wikipedia

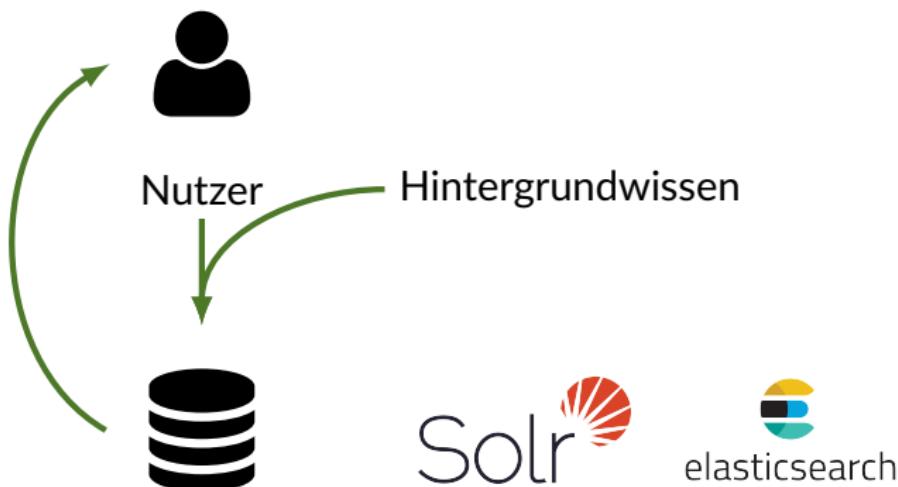
Schema einer Semantischen Suche



Dokumentendatenbank

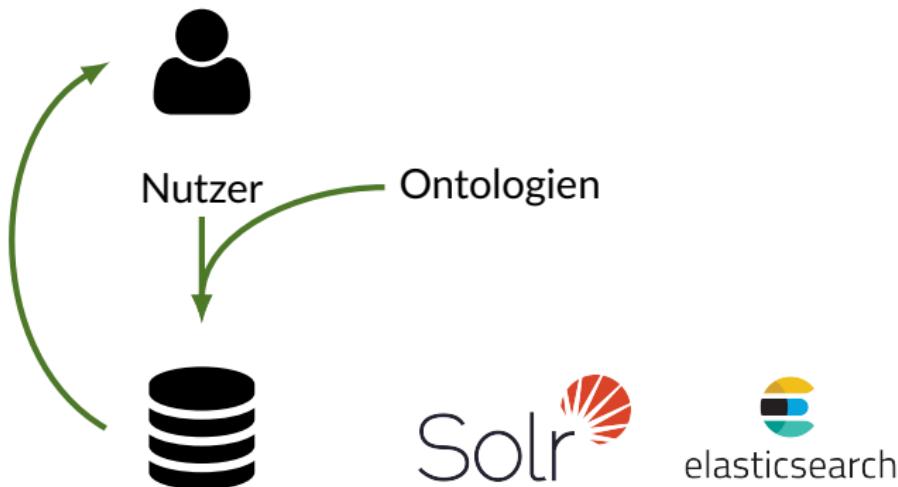


Schema einer Semantischen Suche



Dokumentendatenbank

Schema einer Semantischen Suche



Dokumentendatenbank

Was brauche ich um eine
semantische Suche zu erstellen?

Nutzeranfrage

'Pflanzen mit blauen Blüten'

Nutzeranfrage

'Pflanzen mit blauen Blüten'

TOKENISIERER

Pflanze

mit

blauen

Blüten

Nutzeranfrage

'Pflanzen mit blauen Blüten'

WORTART

Pflanze

mit

blauen

Blüten

NN

ADP

ADJ

NS

Nutzeranfrage

'Pflanzen mit blauen Blüten'

LEMMATISIERER

Pflanze mit blauen Blüten

NN **ADP** **ADJ** **NS**

Pflanze mit blau Blüte

Nutzeranfrage

'Pflanzen mit blauen Blüten'

NAMED ENTITY ERKENNUNG

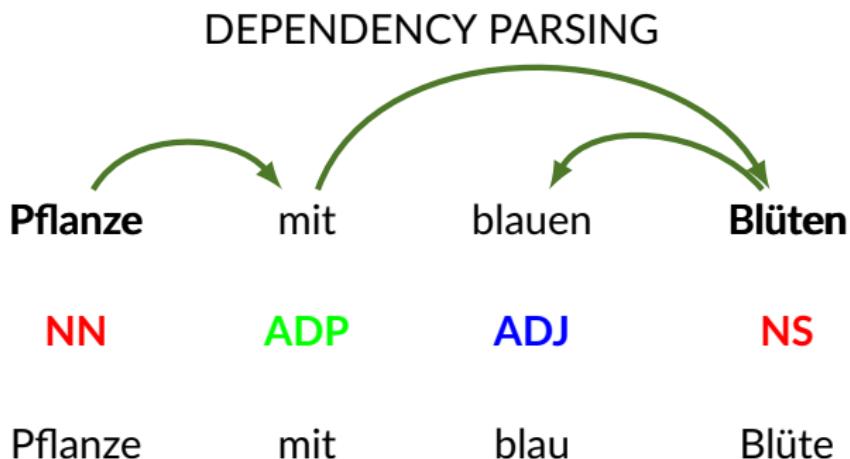
Pflanze mit blauen Blüten

NN ADP ADJ NS

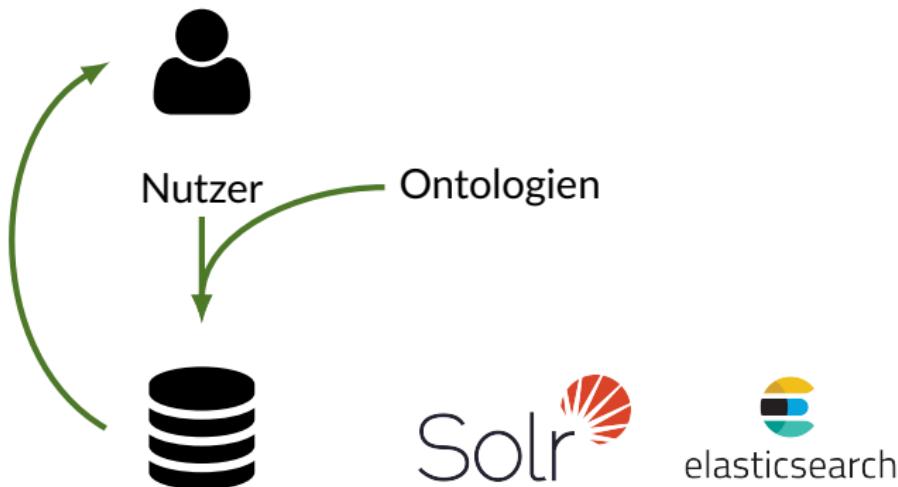
Pflanze mit blau Blüte

Nutzeranfrage

'Pflanzen mit blauen Blüten'

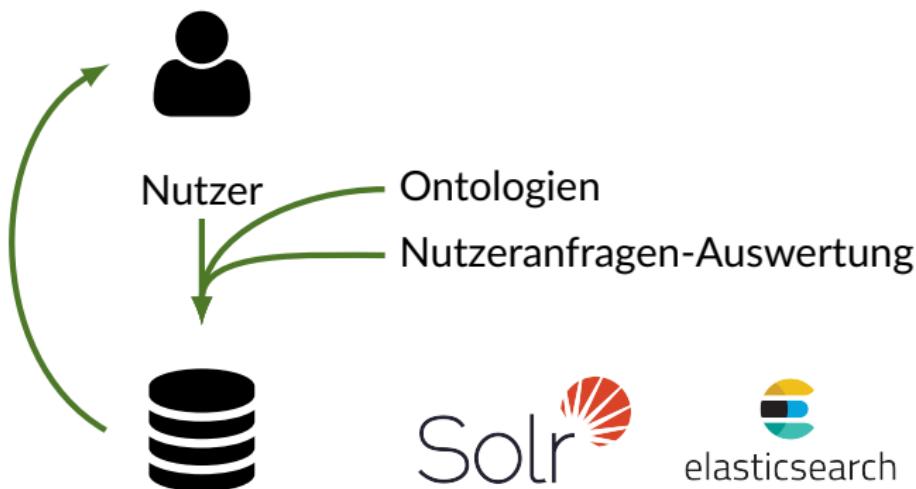


Schema einer Semantischen Suche



Dokumentendatenbank

Schema einer Semantischen Suche

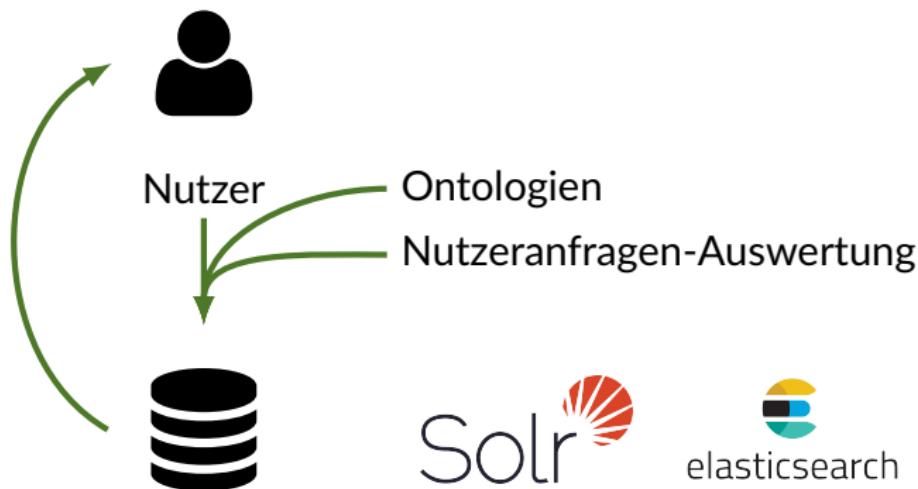


Beispiel eines annotierten Texts

Die **Eibe** kommt als Unterwuchs in dem aus **Fichte** und **Rotbuche** gebildeten Hauptbestand, sowie auch frei in den Felspartien der **Donau** vor. Auf einer Gesamtfläche von 60 ha stehen, nach Schätzung des Forstamts, etwa 600 **Eiben**, d.h. 10 Stück per ha....

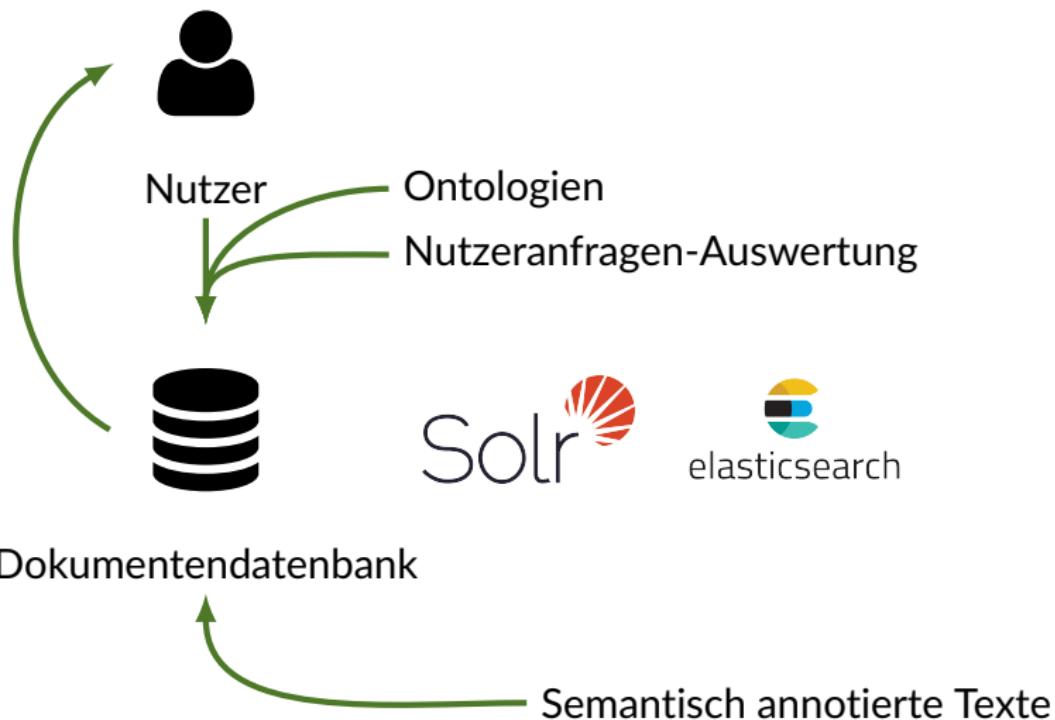
- Klasse des annotierten Wortes ist bekannt
- Bedeutung des annotierten Wortes ist bekannt
- Text wird maschinenlesbar
- Disambiguierung

Schema einer Semantischen Suche



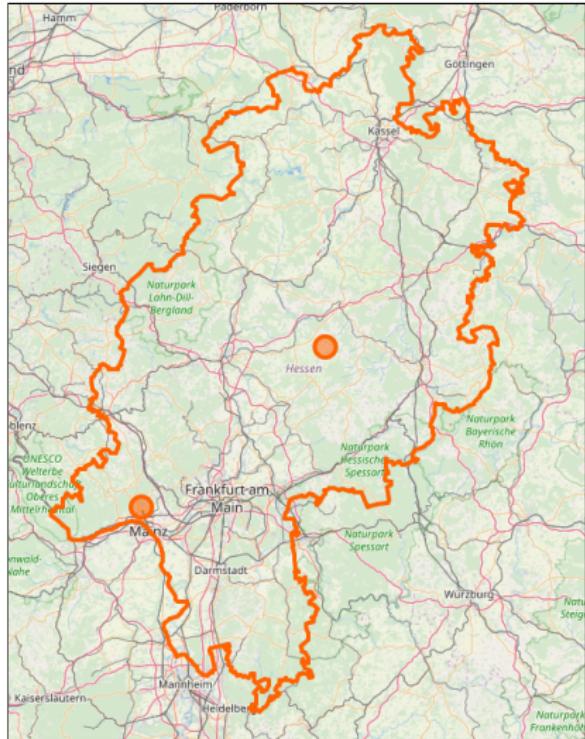
Dokumentendatenbank

Schema einer Semantischen Suche



Geplante Features

- Herunterladen der Suchergebnisse
- Analyse von hochgeladenen Texten
- Zugang auf Textstellen durch Karte
- Erweiterte Merkmals- und Dokumentendatenbank
-
- Einfügen von Geo-Ontologie
- Relation Extraction
- Berechnen von Dokumentvektoren
- Semantische Vektoren von Wörtern in Nutzeranfrage
- Containern der Anwendungen



Hessen in OpenStreetMaps

- Suche nach Ort im Ort/Region
- Ausschließen von Orten

Was bitte ist Relation Extraction?

Die **Eibe** kommt als Unterwuchs in dem aus **Fichte** und **Rotbuche** gebildeten Hauptbestand, sowie auch frei in den Felspartien der **Donau** vor. Auf einer Gesamtfläche von 60 ha stehen, nach Schätzung des Forstamts, etwa 600 **Eiben**, d.h. 10 Stück per ha....

Was bitte ist Relation Extraction?

Die **Eibe** kommt als Unterwuchs in dem aus **Fichte** und **Rotbuche** gebildeten Hauptbestand, sowie auch frei in den Felspartien der **Donau** vor. Auf einer Gesamtfläche von 60 ha stehen, nach Schätzung des Forstamts, etwa 600 **Eiben**, d.h. 10 Stück per ha....

Extrahierte Relationen (Informationen)

- **Eibe** bei **Donau**
- **Eibe** Dichte 10/ha
- Rückreferenz auf Quellsatz

Wie lässt sich die Präzision der Suche weiter verbessern?

- Wortvektoren für einzelne Worte in Nutzeranfrage
- Dokumentvektoren

Wie lässt sich die Präzision der Suche weiter verbessern?

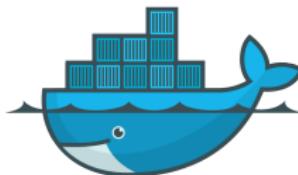
- Wortvektoren für einzelne Worte in Nutzeranfrage
- Dokumentvektoren

Problem

Die Berechnung dieser Vektoren benötigt viele Daten.

Container





Docker Container

- Reproduzierbare Anwendungsumgebung
- Auslieferung der kompletten Anwendung an den Nutzer
- Nutzung ist nur eingeschränkt durch die Hardware

Dankeschön

Vielen Dank!