

Seminar Texttechnologien und Semantic Web

Ontologie-basiertes Information sharing und Ontologie merging

22. Juni 2005

Marian Dörk

Übersicht

1. Einleitung
2. Heterogenität
3. Definition
4. Anwendung
5. Umsetzung
6. Ausblick

Einleitung

Information

- ✓ erzeugen, speichern und transportieren
- ★ finden und integrieren = *Information Sharing*

Semantik von Information lesbar für Computer

Heterogenität

zwischen verschiedenen Informationsquellen

Syntax

Struktur

Semantik

Syntax

Muster und Regeln für Satzbau

verschiedene Dateiformate

- ▶ mit XML kein Problem mehr

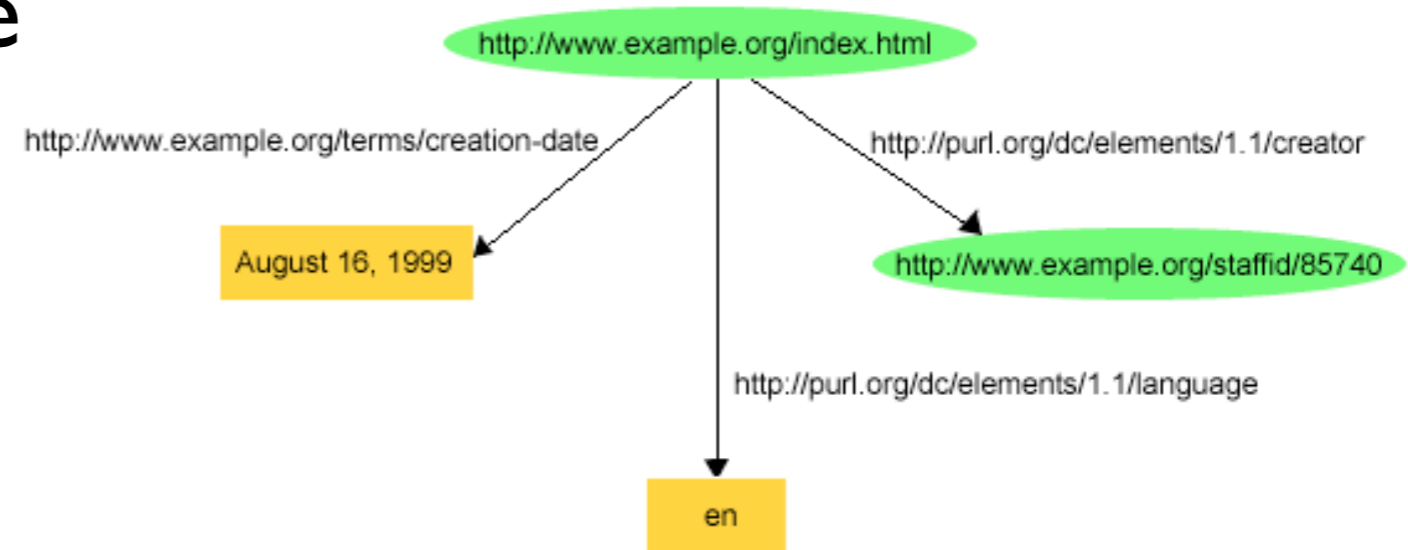


Struktur

konzeptueller Aufbau einer Quelle

potenzielle Konflikte

- Namensgebung
- Datentypen
- fehlende Attribute



► Resource Description Framework (RDF)

Semantik

Sinn und Bedeutung der Quelle

Erfassung: Text, Struktur, explizit

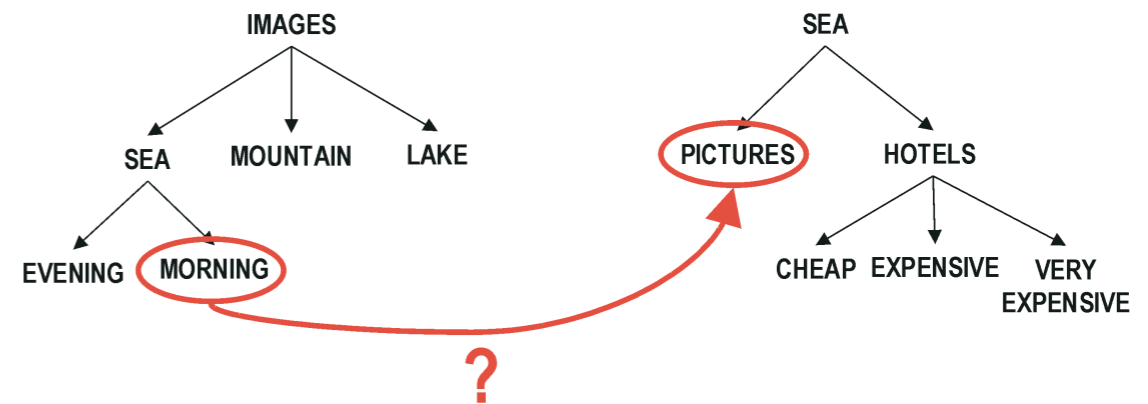
potenzielle Konflikte

- Daten: unterschiedliche Skalen und Wertebereiche
- Domäne: Einordnung, Überlappung, Inkonsistenz

Semantik

Semantische Interoperabilität

- Namen und Bezeichnungen: Synonymmengen
- Begriffsnetzwerke
- Konzeptgitter
- Merkmale und Bedingungen



Definition

Philosophie

Die Ontologie beschäftigt sich (primär) mit dem Sein, dem Seienden als solchem und mit den fundamentalen Typen von Entitäten.

Informatik

Unter einer Ontologie versteht man [...] ein formal definiertes System von Begriffen und/oder Konzepten und Relationen zwischen diesen Begriffen.

Definition

“An ontology is an explicit specification of a conceptualization.”

Gruber, 1993

explizite Repräsentation von Semantik

- geteiltes Vokabular
- Konzeptualisierung einer Thematik
- ▶ zugänglich für Computer

Anwendung

Kommunikation

Entwicklung komplexer Systeme

Interoperabilität

Informationsgewinnung

Kommunikation

Information Communities

Terminologien erklären (übersetzen)

Beispiel: Entity-Relationship-Diagramme

Entwicklung

Planung und Umsetzung komplexer Systeme

Wiederverwendung existierender Software

Beispiel: API-Referenzen

Interoperabilität

Integration existierender Systeme

Ontologien als 'interlinguas'

Beispiel: WSDL für Web Services

Informationsgewinnung

Mehrdeutigkeiten bei einfacher Textsuche

Nutzeranfrage kann Synonyme einschließen

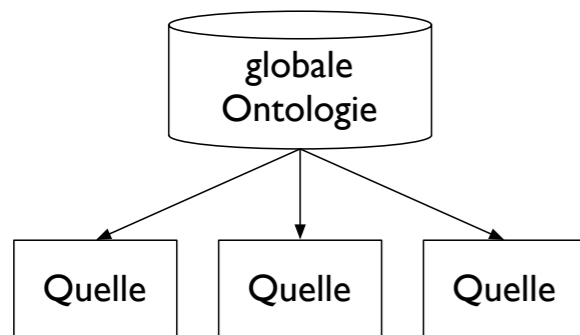
Beispiel: Verwendung von WordNet

Umsetzung

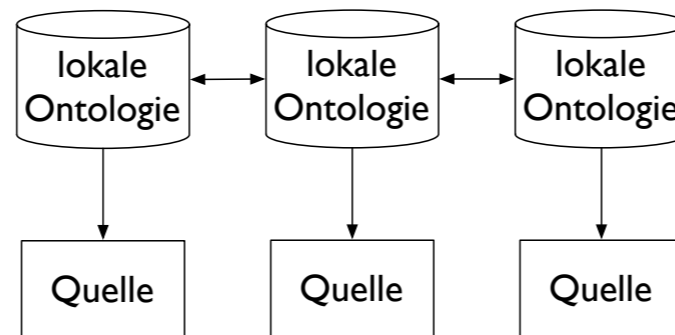
Herangehensweisen

Infrastruktur für Information Sharing

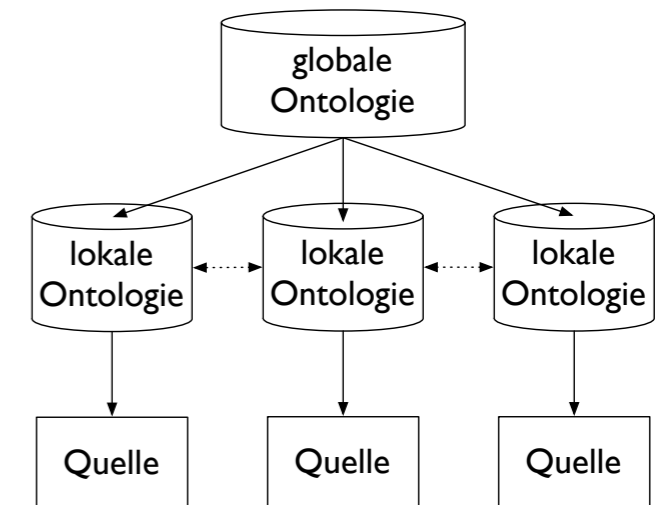
Herangehensweisen



single-ontology



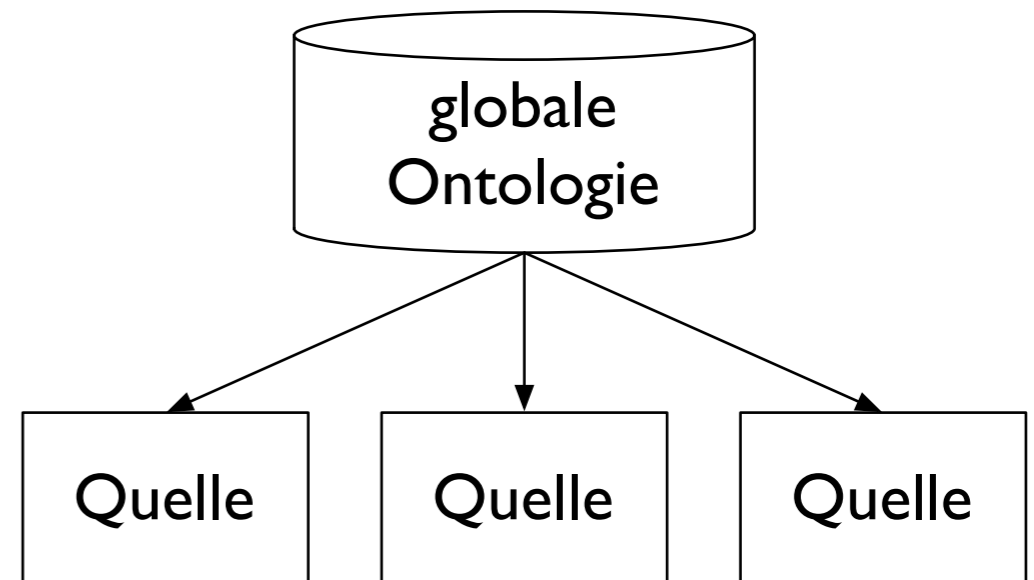
multiple ontologies



hybrid approach

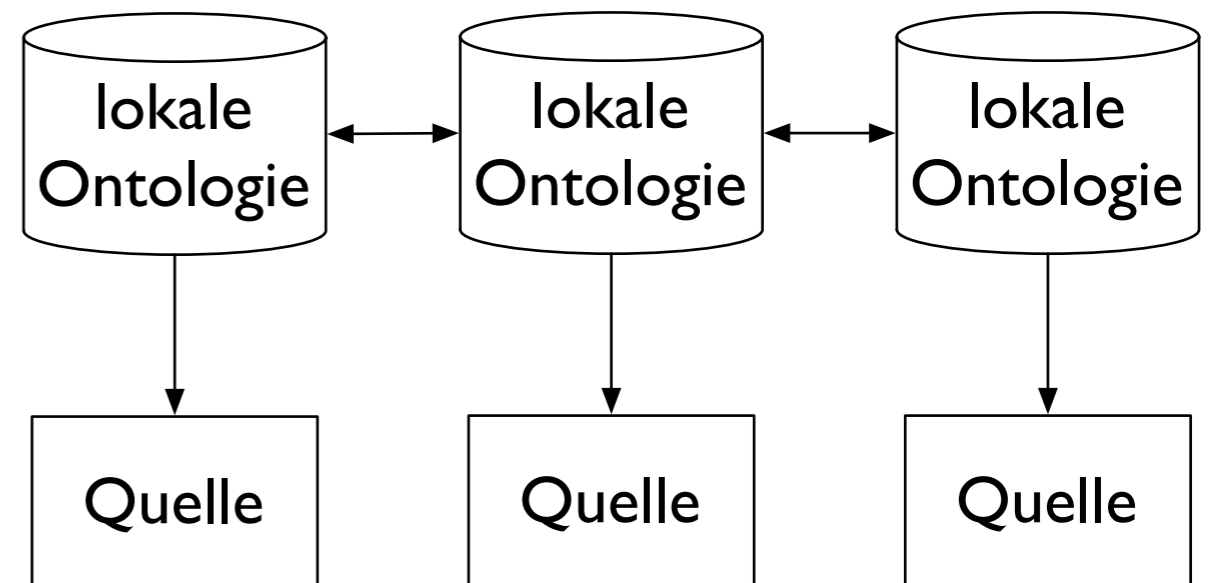
single-ontology

Hinzufügen neuer
Quellen sehr komplex



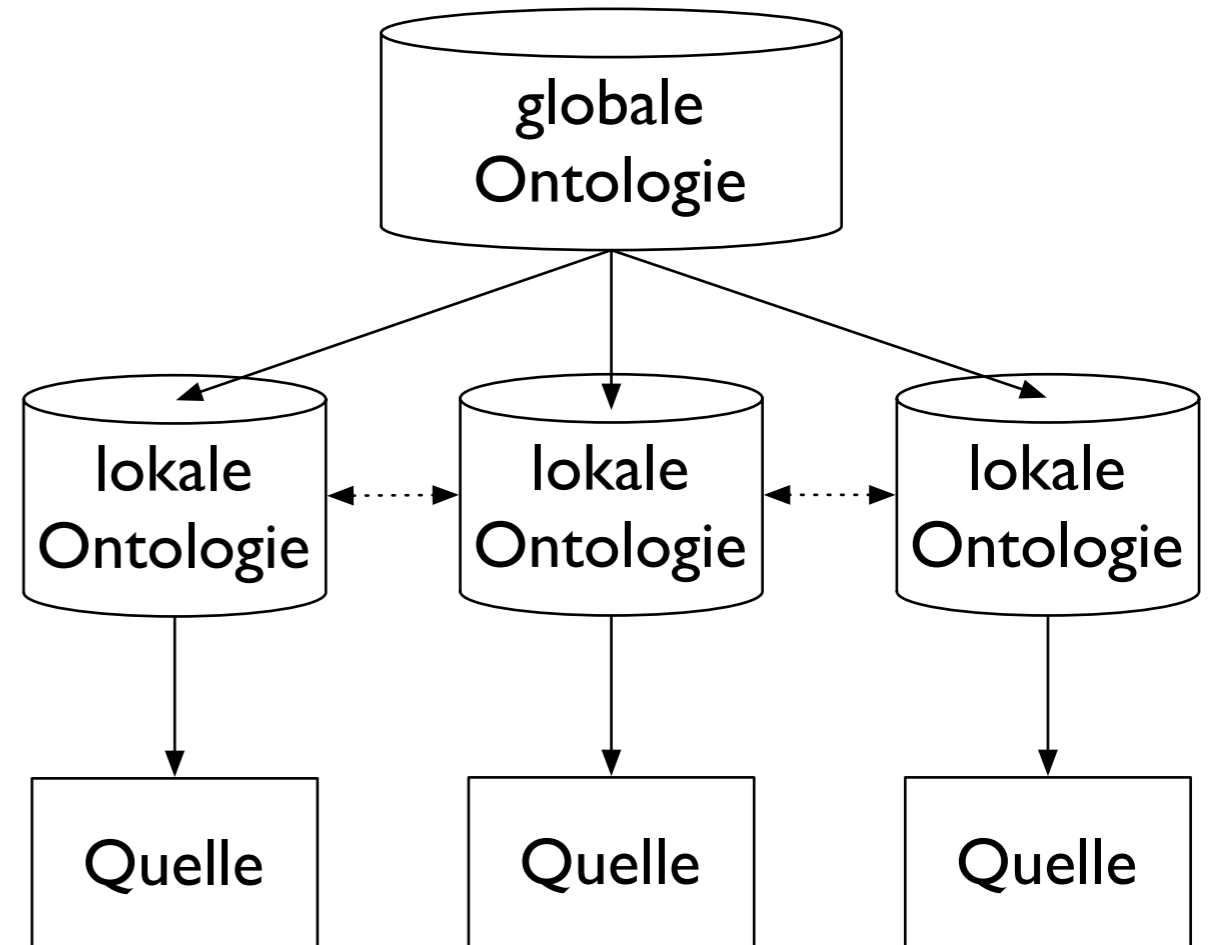
multiple ontologies

Integration und
Vergleich sehr
komplex

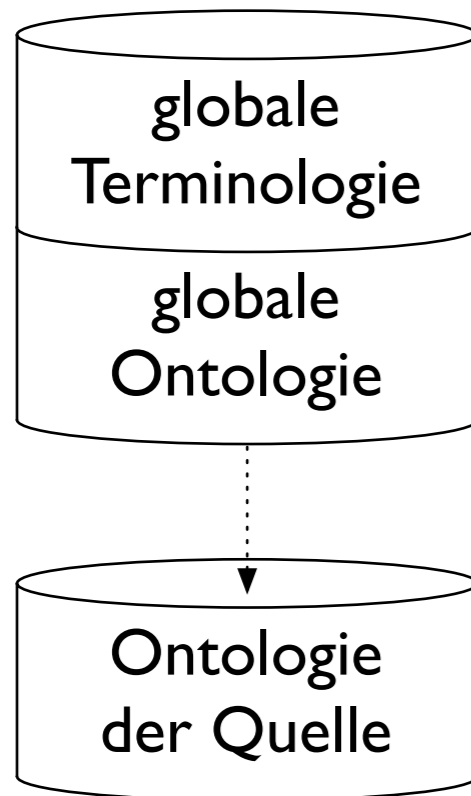


hybrid approach

Hinzufügen und
Abfragen einfacher



Infrastruktur



globale Terminologie $\{W, I\}$

$W \sim$ Menge von Wörtern

$I : W \times W \rightarrow \{\text{syno, hyper, hypo}\}$

globale Ontologie $\{ST, T, R, A\}$

$ST = \{W_L, I_L\}$

$T \sim$ Menge an Begriffen aus W_L

$R \subseteq T \times T \sim$ Relationen

$A \sim$ Axiome in der Form: $T_i \sqsubseteq T_j$

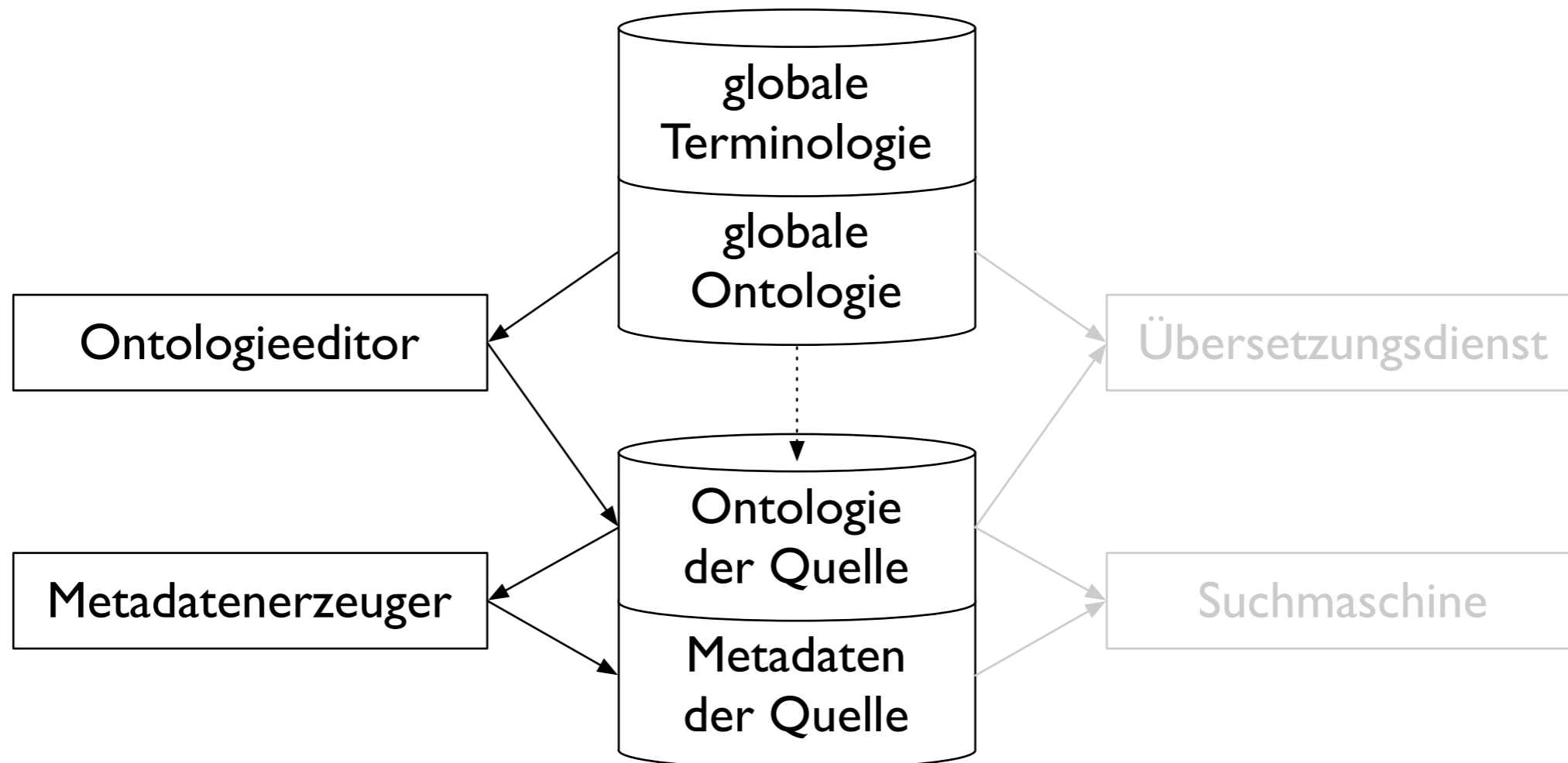
Ontologie der Quelle $\{S, C, d\}$

$S = \{ST_s, T_s, R_s, A_s\}$

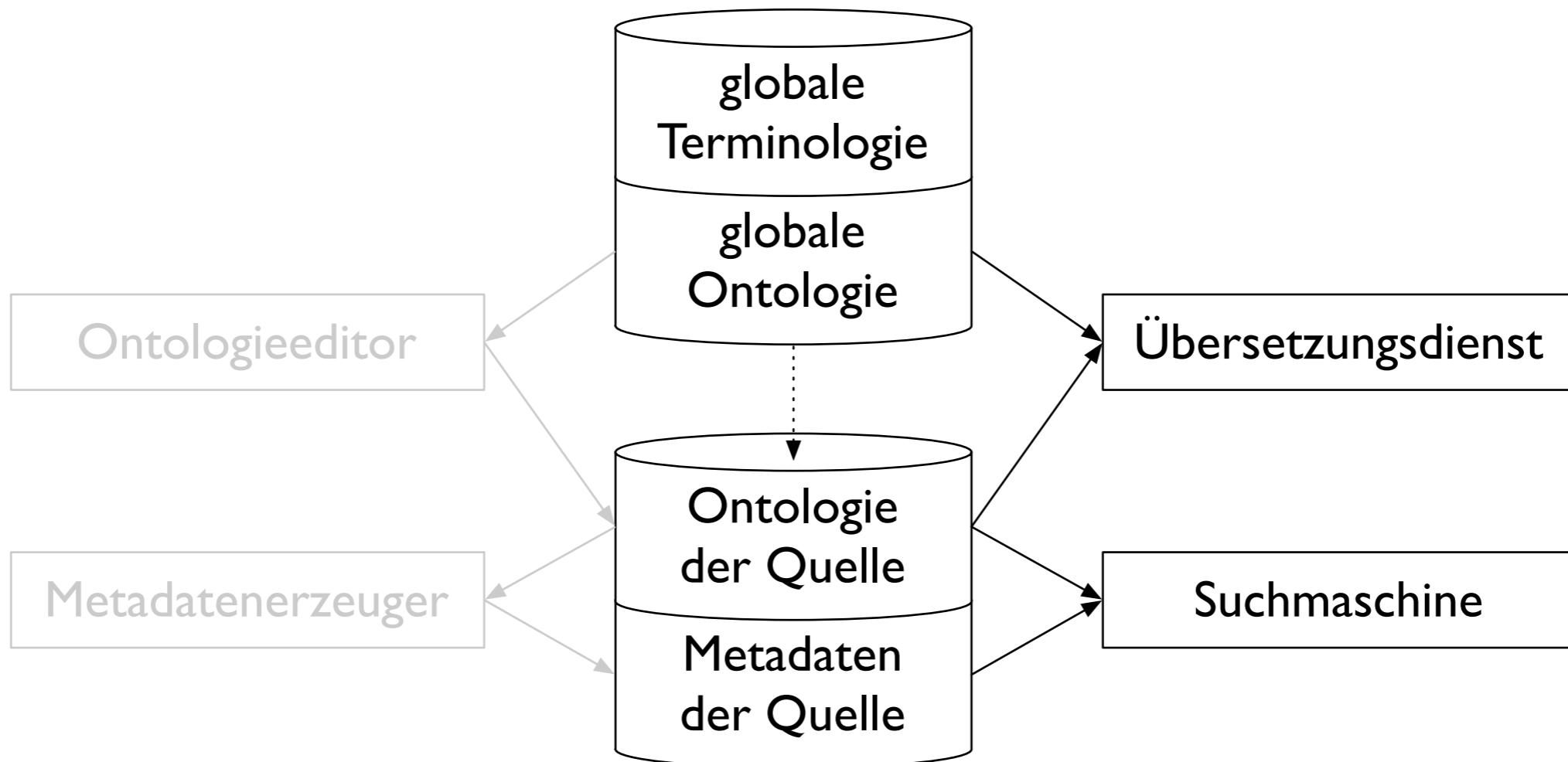
$C \sim$ Menge an Klassennamen nicht in S

$d \sim$ Definitionen von C_i mit Hilfe von R_s

Entwicklung und Pflege



Abfragen und Integration



Ausblick

Forschung schon sehr weit

Praxis noch am Anfang

Erstellung lokaler und globaler Ontologien sehr aufwändig

Danke für eure Aufmerksamkeit