

D3.4 Softwarekomponenten





]init[

Inhaltverzeichnis

1	Einleitung		3
2	Editortools		4
	2.1	Einleitung	4
	2.2	VocBench	4
	2.3	Protégé	7
	2.4	WebProtégé	11
	2.5	Entryscape	14
	2.6	Erklärung der Tabellenkategorien	16
3	Visualisierungstools		20
	3.1	Einleitung	20
	3.2	ShowVoc	20
	3.3	Stardog Explorer	23
	3.4	Sparnatural	25
	3.5	OWLGrEd	28
	3.6	WebVOWL	30
	3.7	Erklärung der Tabellenkategorien	33
4	Tripelstores		36
	4.1	Einleitung	36
	4.2	Stardog	37
	4.3	Graph DB	38
	4.4	RDF4J	39
	4.5	TypeDB	40
5	Fazit	und Ausblick	41





1 Einleitung

Im Rahmen des RegCheck Projektes wurde die potenzielle Rolle die Wissensgraphen und Semantic Web im Rahmen der Registermodernisierung spielen können, erforscht.

Die Entwicklung und Integration von Wissensgraphen erforderten den Einsatz spezialisierter Werkzeuge, die die Modellierung, Visualisierung und Verwaltung komplexer Wissensstrukturen ermöglichen. Um perspektivisch eine dauerhafte Implementierung des entwickelten Wissensgraphen in zukünftigen Anwendungen vorzubereiten, ist es daher von entscheidender Bedeutung, eine erste Evaluation verschiedener Tools vorzunehmen, die für eine Implementierung relevant wären.

Dieses Dokument beschreibt untersuchte Werkzeuge aus drei wesentlichen Kategorien:

- 1. Editiertools für die Wissensgraph-Entwicklung und -Bearbeitung.
- 2. Visualisierungstools zur Darstellung und Analyse von Wissensgraphen.
- 3. Tripelstores für die effiziente Speicherung und Verwaltung des Wissensgraphen.

Die Auswahl der evaluierten Werkzeuge erfolgte auf der Grundlage mehrerer Kriterien: Zum einen wurde berücksichtigt, wie häufig die Werkzeuge in wissenschaftlichen Konferenzen und Forschungspublikationen genannt werden, da dies ein Indikator für ihre Relevanz und Akzeptanz in der Fachgemeinschaft ist. Zum anderen wurde ihre allgemeine Bekanntheit und praktische Verbreitung herangezogen, um sicherzustellen, dass die ausgewählten Werkzeuge bereits in einer Vielzahl von Anwendungsszenarien erfolgreich eingesetzt wurden und sich als zuverlässig und effizient erwiesen haben.

Ziel dieses Evaluationsprozesses war es, fundierte Erkenntnisse über die für eine mögliche langfristige Implementierung des Wissensgraphen erforderlichen Werkzeuge zu gewinnen. Die Untersuchung sollte aufzeigen, welche Werkzeuge die Anforderungen an eine robuste und skalierbare Wissensgraph-Implementierung am besten erfüllen können und welche Potenziale und Herausforderungen mit ihrem Einsatz verbunden sind. Es wurden sowohl funktionale als auch nicht-funktionale Aspekte der Werkzeuge berücksichtigt, um ein ganzheitliches Bild ihrer Leistungsfähigkeit zu zeichnen. Um eine möglichst gute Vergleichbarkeit innerhalb der einzelnen Kategorien zu erhalten, wurde für jede der Kategorien, sofern möglich, eine eigenes Bewertungsschema entwickelt, welches auf die verschiedenen Werkzeuge angewendet wurde.

Diese Analyse soll als Grundlage für eine spätere Entscheidung dienen, welche Werkzeuge im Rahmen einer Implementierung des Wissensgraphen außerhalb des aktuellen experimentellen Umfelds, benötigt werden. Dazu stellt diese Analyse toolübergreifende Auswahlkriterien auf.





2 Editortools

2.1 Einleitung

Editortools spielen eine wichtige Rolle im Prozess der Entwicklung und Gestaltung von Wissensgraphen. Diese spezialisierten Softwarewerkzeuge ermöglichen es, komplexe semantische Strukturen zu modellieren, zu bearbeiten und zu verwalten, indem sie eine benutzerfreundliche Schnittstelle bieten, die die Nutzenden bei der Erstellung und Pflege ontologischer Modelle und den daraus abgeleiteten Wissensgraphen unterstützt. Durch die Bereitstellung von Funktionen zur Definition von Klassen, Relationen und Eigenschaften sowie zur Organisation dieser Elemente in einem strukturierten und konsistenten Arbeitsframework tragen Editortools wesentlich zur nutzerfreundlichen Bearbeitung von Wissensgraphen bei.

Das RegCheck Modell wurde mithilfe "klassischer" Erstellungs- und Bearbeitungsmethoden einer Ontologie erstellt. Dabei werden Tripel direkt in Quelltext-Editoren geschrieben und bearbeitet.

Im Vergleich dazu bieten Ontologie-Editortools eine Möglichkeit, durch visuelle und interaktive Funktionen die Modellierung und Verwaltung semantischer Strukturen wesentlich zu vereinfachen. Diese Tools minimieren die Fehleranfälligkeit und verbessern die Effizienz des Entwicklungsprozesses erheblich, indem sie eine weniger abstrakte Herangehensweise an die Ontologieerstellung ermöglichen. Dadurch werden Hürden für Verständlichkeit und Wartbarkeit des Wissensgraphen verringert. Dies ist insbesondere relevant für eine spätere Pflege und Wartung des Wissensgraphen.

Ein weiterer bedeutender Aspekt ist die Integration in größere semantische Ökosysteme. Die Tools ermöglichen den Import und Export von Ontologien in verschiedenen Formaten, was die Zusammenarbeit und Wiederverwendung erleichtert. Sie bieten die Werkzeuge, um komplexes Wissen formal zu modellieren und zu verwalten, und tragen dazu bei, dass die Bearbeitung von Wissensgraphen für eine größere Anzahl an Personen möglich ist.

2.2 VocBench

VocBench ist ein vom Publication Office der Europäischen Union herausgegebenes Editiertool für Vokabulare, Thesauri und Ontologien. Es ist eine Open Source Lösung und hat als Zielgruppe in erster Linie EU-Institutionen, kann jedoch auch von anderen Akteuren des öffentlichen Sektors verwendet werden.

Es kann sowohl auf eigenen Servern als auch auf den Servern des Publication Office genutzt werden. Nur im Falle von eigenem hat man die Möglichkeit Projektumgebungen selbständig anzulegen. Sofern man die Server des Publication Office nutzt, müssen dort alle Projektumgebungen voreingestellt und freigegeben werden.





VocBench umfasst alle grundlegenden Funktionalitäten eines Editiertools und weist hierbei keine besonderen Vor- oder Nachteile auf. Eine herausragende Eigenschaft sind bereits hinterlegte Vokabulare und Ontologien des Publication Office, die bei Bedarf der eigenen Projektumgebung hinzugefügt werden können. So ist es möglich, auf das zusätzliche Importieren dieser zu verzichten. Jedoch zeichnet sich VocBench durch eine geringere Benutzerfreundlichkeit aus (in der Tabelle dargestellt durch eine 3-4 Bewertung). Die in Teilen komplizierte Dokumentation, kombiniert mit einer vergleichsweise kleinen Nutzergemeinschaft, welche zur Unterstützung herangezogen werden kann, erschwert sowohl den Einstieg als auch den Umgang mit eventuell auftretenden Problemen bei VocBench. Personen, die mit VocBench arbeiten möchten, sollten bereits tiefergehende Kenntnisse in der Ontologieerstellung aufweisen.

Zusammenfassend beschrieben, ist VocBench ein grundsätzlich umfangreiches Werkzeug, welches jedoch hohe Einstiegs- und teilweise auch Benutzungshürden ausweist.

Tabellarische Bewertung VocBench

Anbieter	Publications Office der Europäischen Union
Land	EU
Anbieterwebsite	https://op.europa.eu/de/web/eu-vocabularies/vocbench
Kosten	Beim Nutzen der VocBench der EU fallen keine Kosten an, beim Hosten in Eigenregie fallen Kosten für die Installation, den Betrieb und Wartung an.
Kollaborationsmöglichkeit	Verschiedene Nutzenden können es innerhalb von VocBench Webanwendung parallel bearbeiten.
Fähigkeitslevel nach Fä- higkeitsskala	3-4
Liste editierbarer Ontolo- giesprachen	RDF, OWL, SKOS, OntoLex, EDOAL
Bearbeitungsmöglichkeit	"What you see, is what you get" Bearbeitung
SHACL möglich	Nein





Unterstützung beim Schreiben von SHACL	Nein
Unterstützung beim Aus- führen von SHACL	Nein
Unterstützung beim De- buggen von SHACL	Nein
SPARQL möglich	Ja
Unterstützung beim Schreiben von SPARQL- Abfragen	Ja. Ein Syntax completion helper, der basierend auf den eingegebenen Werten Vorschläge macht. Basiert auf dem Open Source YASGUI SPARQL Editor.
Unterstützung beim Ausführen von SPARQL-Abfragen	Ja
Unterstützung beim Debuggen von SPARQL-Abfragen	Nein
Offizielle Dokumentation vorhanden	Ja https://vocbench.uniroma2.it/doc/
Community-Support	Es gibt eine <u>VocBench Nutzenden Gruppe</u> auf Google Groups. Nicht alle Fragen werden beantwortet, evtl. auf- grund der relativ kleinen Gruppe an Nutzenden.
Ist die Software frei ver- fügbar?	Ja
Lizenz	BSD 3-Clause
Import Möglichkeiten	BinaryRDF, JSON-LD, N-Quads, N-Triples, N3, NDJSON-LD, RDF/JSON, RDF/XML, TriG, TriG-star, TriX, Turtle, Turtle-star





	Es kann sowohl aus Dateien als auch aus Triplestores importiert werden
Export Möglichkeiten	BinaryRDF, JSON-LD, N-Quads, N-Triples, N3, NDJSON-LD, RDF/JSON, RDF/XML, TriG, TriG-star, TriX, Turtle, Turtle-star
Vorab beinhaltete Taxo- nomien/ Ontologien	OWL und RDF-Klassen und Properties, bestehende EU-Ontologien wie z.B. Digital Europa Thesaurus, Europass tables, ESCO taxonomies sind bereits in VocBench vorhanden und können bei Bedarf in das Projekt hinzugefügt werden
Editierbarer Namespace: IRI/ URN etc.	Ja
4-Augen Prinzip o.Ä. mög- lich	Ja, VocBench bietet eine optionale Funktion an, mit der jede Änderung noch einmal extra bestätigt werden muss
Installationsort: Web oder Lokal?	Beides möglich: Einmal in Form einer Webanwendung über das Publications Office und einmal auf eigenen Servern
Oberflächensprache(n)	Deutsch, Englisch, Spanisch, Französisch, Italienisch
PlugIns/ Anpassungsmög- lichkeiten	Keine, jedoch können beim Publications Office Änderungen/ Anpassungen beantragt bzw. vorgeschlagen werden
Sonstiges	

2.3 Protégé

Protégé Desktop, eine Open-Source-Software der Stanford University, ist eines der am weitesten verbreitete Werkzeuge für die Ontologieentwicklung. Es hat sich insbesondere im akademischen Bereich etabliert.

Ein wichtiges Merkmal von Protégé Desktop ist seine umfassende Dokumentation, die einen niedrigschwelligen Zugang zur Ontologieentwicklung ermöglicht. Die verfügbare Literatur und die Online-Community bieten eine Fülle an Ressourcen, die Anwendern dabei helfen, sich schnell mit der Software vertraut zu machen.





Jedoch ist Protégé insbesondere im Hinblick auf seine Benutzeroberfläche veraltet. Obwohl das Interface funktional und intuitiv ist, entspricht es nicht mehr den modernen Designstandards, was trotz guter Dokumentation den Einstieg erschwert.

Eine der herausragendsten Eigenschaften von Protégé Desktop ist die Vielzahl von Plugins, welche es erlauben zusätzliche Funktionen und Anpassungen zu integrieren. Diese Plugins erweitern die Kernfunktionalitäten der Software und bieten Lösungen für spezialisierte Anforderungen. Von der semantischen Abfrage bis hin zu spezifischen Visualisierungen bieten die Plugins eine beachtliche Flexibilität, die Protégé Desktop von anderen Ontologie-Editoren abhebt.

Jedoch stellt die mangelnde Aktualität vieler dieser Plugins eine erhebliche Herausforderung dar. Viele der ursprünglich entwickelten Erweiterungen werden nicht mehr aktiv gepflegt, was zu Kompatibilitätsproblemen und funktionalen Einschränkungen führt. Die fehlende Wartung und Weiterentwicklung dieser Plugins mindert den Nutzen und die Effizienz von Protégé Desktop, insbesondere wenn Anwender auf spezifische, nicht mehr unterstützte Funktionen wie z.B. SHACL angewiesen sind. Dies stellt ein Problem dar, das den ansonsten positiven Gesamteindruck des Tools trübt. Zugleich ist zu hoffen, dass in Zeiten steigenden Interesses an Wissensgraphen einige der PlugIns aktualisiert und angepasst werden.

Tabellarische Bewertung Protégé Desktop

Anbieter	Standford University
Land	USA
Anbieterwebsite	https://protege.stanford.edu/
Kosten	Protégé: o€
Kollaborationsmöglichkeit	Asynchron über Drittanbieterlösungen möglich
Fähigkeitslevel nach Fä- higkeitsskala	2
Liste editierbarer Ontolo-	RDF/XML
giesprachen	Turtle
	OWL/XML
	Manchester OWL Syntax
	Functional OWL Syntax





Bearbeitungsmöglichkeit	"What you see, is what you get" Bearbeitung
SHACL möglich	Ja, mithilfe von Plugins
Unterstützung beim Schreiben von SHACL	Abhängig von Plugin
Unterstützung beim Ausführen von SHACL	Abhängig von Plugin
Unterstützung beim De- buggen von SHACL	Abhängig von Plugin
SPARQL möglich	Ja, mithilfe von Plugins
Unterstützung beim Schreiben von SPARQL- Abfragen	Abhängig von Plugin
Unterstützung beim Ausführen von SPARQL-Abfragen	Abhängig von Plugin
Unterstützung beim Debuggen von SPARQL-Abfragen	Nein
Offizielle Dokumentation vorhanden	Ja http://protegeproject.github.io/protege/
Community-Support	Ja, in Form einer User Support Mailing List https://protege.stanford.edu/support.php
Ist die Software frei ver- fügbar?	Ja
Lizenz	BSD 2 Clause
Import Möglichkeiten	RDF/XML





linit

Turtle

OWL/XML

Manchester OWL Syntax

Functional OWL Syntax

Export Möglichkeiten RDF/XML

Turtle

OWL/XML

Manchester OWL Syntax

Functional OWL Syntax

Vorab beinhaltete Taxonomien/ Ontologien Nein

Editierbarer Namespace:

IRI/ URN etc.

Ja

4-Augen Prinzip o.Ä. mög-

lich

Nein

Installationsort: Web oder

Lokal?

Lokal

Oberflächensprache(n) E

Englisch

PlugIns/ Anpassungsmög-

lichkeiten

Durch eine große Community sind viele Plugins entwickelt

worden, welche das native Angebot erweitern.

Sonstiges Eine Vielzahl an PlugIns, aber viele seit 3+ Jahren nicht mehr

aktualisiert.





2.4 WebProtégé

WebProtégé ist die browserbasierte Variante von Protégé. Als eine moderne Ergänzung zur etablierten Desktop-Version bietet WebProtégé eine benutzerfreundlichere Plattform zur Erstellung, Verwaltung und gemeinsamen Bearbeitung von Ontologien direkt im Browser.

Ein entscheidender Vorteil von WebProtégé liegt in seiner modernen und intuitiven Benutzeroberfläche. Im Gegensatz zu Protégé wurde WebProtégé so gestaltet, dass es den aktuellen Standards von Webanwendungen entspricht. Das Design ist klar strukturiert und ermöglicht eine deutlich flüssigere und effizientere Interaktion mit dem System. Diese moderne Benutzeroberfläche macht WebProtégé besonders attraktiv für Anwender ohne tiefgehende Vorerfahrungen. Alle grundlegenden Funktionalitäten und die Logik einer Ontologie erschließen sich bereits nach kurzer Zeit und sind in allgemeingültiger Sprache verfasst.

Obwohl WebProtégé den Einstieg in die Ontologieerstellung vereinfacht, bietet es im Vergleich zu Protégé einen reduzierten Funktionsumfang. Die Flexibilität und Anpassbarkeit sind in WebProtégé vergleichsweise eingeschränkt. Dies erhöht die Benutzerfreundlichkeit, jedoch kann diese wiederum dazu führen, dass bei komplexeren Projekten Grenzen auftreten. Das Auftreten solcher Limitationen sind jedoch vermehrt in sehr komplexen akademischen Projekten mit speziellen Anforderungen zu erwarten, nicht in Modellen mit einer begrenzten Anzahl von Elementarten wie dem RegCheck Modell.

Tabellarische Bewertung Webprotégé

Anbieter	Standford University
Land	USA
Anbieterwebsite	https://protege.stanford.edu/
Kosten	Keine
Kollaborationsmöglich- keit	Es kann innerhalb des Tools ein Projekt mit anderen Nutzenden geteilt und dann kollaborativ bearbeitet werden
Fähigkeitslevel nach Fä- higkeitsskala	1
Liste editierbarer Ontolo-	RDF/XML
giesprachen	Turtle
	OWL/XML





Manchester OWL Syntax
Functional OWL Syntax

	T Offictional OWE Syntax
Bearbeitungsmöglichkeit	"What you see, is what you get" Bearbeitung
SHACL möglich	Nein, nur in Desktop Variante
Unterstützung beim Schreiben von SHACL	Nein, nur in Desktop Variante
Unterstützung beim Aus- führen von SHACL	Nein, nur in Desktop Variante
Unterstützung beim De- buggen von SHACL	Nein, nur in Desktop Variante
SPARQL möglich	Nein aber ein UI mit gleicher Nutzerfunktionalität ist möglich
Unterstützung beim Schreiben von SPARQL- Abfragen	Nein, aber äquivalente UI
Unterstützung beim Aus- führen von SPARQL-Ab- fragen	Nein, aber äquivalente UI
Unterstützung beim De- buggen von SPARQL-Ab- fragen	Nein
Offizielle Dokumentation vorhanden	Ja https://protegewiki.stanford.edu/wiki/WebProte-geUsersGuide
Community-Support	Ja, in Form einer User Support Mailing List https://protege.stanford.edu/support.php
Ist die Software frei ver- fügbar?	Ja





Lizenz	BSD 2 Clause
Import Möglichkeiten	RDF/XML
	Turtle
	OWL/XML
	Manchester OWL Syntax
	Functional OWL Syntax
Export Möglichkeiten	RDF/XML
	Turtle
	OWL/XML
	Manchester OWL Syntax
	Functional OWL Syntax
Vorab beinhaltete Taxo- nomien/ Ontologien	Nein
Editierbarer Namespace: IRI/ URN etc.	Ja
4-Augen Prinzip o.Ä. möglich	Nein
Installationsort: Web oder Lokal?	Web
Oberflächensprache(n)	Englisch
Plugins/ Anpassungs- möglichkeiten	Keine für Webprotégé, aber durch eine große Community sind viele Plugins entwickelt worden, welche das Angebot welches Protégé (Desktop) nativ erlaubt, erweitert.
Sonstiges	





2.5 Entryscape

Entryscape ist ein Ontologie-Editor, der vergleichsweise neu auf dem Markt ist und noch aktiv weiterentwickelt wird. Die Ansätze sind vielversprechend, wenn auch noch nicht vollständig ausgereift. Auch sind einige Funktionalitäten wie z.B. SHACL Unterstützung noch nicht umfassend implementiert. Dies schränkt die Aussagekraft über das volle Potenzial und die Leistungsfähigkeit des Tools ein. Dieser Entwicklungsstand bedeutet, dass es noch einige Unsicherheiten gibt, wie effektiv das Tool in der Praxis eingesetzt werden kann, insbesondere im Bereich der Ontologiebearbeitung.

Trotz dieser Einschränkungen wurde Entryscape in dieser Analyse berücksichtigt, vor allem weil es sich um eine neue Entwicklung handelt. Der Markt für Editiertools ist von älteren Systemen geprägt, die oft nicht mehr den aktuellen Anforderungen und technologischen Standards entsprechen. Entryscape bietet hier die Möglichkeit, von den neuesten technologischen Fortschritten zu profitieren.

Die Init AG vertestet Entryscape aktuell und in naher Zukunft auch in anderen Projekten, wodurch wertvolle Einblicke in die Weiterentwicklung gewonnen werden können. Diese laufende Nutzung ermöglicht es uns, zeitnah zu beobachten, ob sich Entryscape in Zukunft zu einem nützlichen Werkzeug für die Ontologiebearbeitung entwickelt. Zum jetzigen Zeitpunkt ist es noch nicht dazu geeignet, besitzt aber das Potential dazu, sodass eine erneute Betrachtung in Zukunft empfehlenswert ist.

Tabellarische Bewertung Entryscape

Anbieter	Entryscape
Land	Schweden
Anbieterwebsite	https://entryscape.com/en/
Kosten	450€ monatlich für das kleinste Paket
Kollaborationsmöglichkeit	Mehrere Personen können auf ein Projekt zugreifen, gleichzeitiges arbeiten ist möglich. Es ist aber noch nicht zu 100% zuverlässig und weist Schwierigkeiten auf
Fähigkeitslevel nach Fä- higkeitsskala	3





Liste editierbarer Ontolo- giesprachen	Keine Übersicht vorhanden
Bearbeitungsmöglichkeit	"What you see, is what you get" Bearbeitung
SHACL möglich	Nein, jedoch vom Entwickler angekündigt
Unterstützung beim Schreiben von SHACL	Nein
Unterstützung beim Aus- führen von SHACL	Nein
Unterstützung beim De- buggen von SHACL	Nein
SPARQL möglich	Nein
Unterstützung beim Schreiben von SPARQL- Abfragen	Nein
Unterstützung beim Aus- führen von SPARQL-Ab- fragen	Nein
Unterstützung beim De- buggen von SPARQL-Ab- fragen	Nein
Offizielle Dokumentation	Ja
vorhanden	ENTRYSCAPE Dokumentation
Community-Support	Nein
Ist die Software frei ver- fügbar?	Free Version ist mit eingeschränkten Funktionen verfügbar
Lizenz	Small Package kostet 450€, weitere Kosten auf Nachfrage





Import Möglichkeiten	RDF/ XML Dateien, keine weitere Spezifikation
Export Möglichkeiten	Unklar
Vorab beinhaltete Taxo- nomien/ Ontologien	Dublin Core Terms, SKOS
Editierbarer Namespace: IRI/ URN etc.	Ja, kann angepasst werden
4-Augen Prinzip o.Ä. mög- lich	Nein
Installationsort: Web oder Lokal?	Web
Oberflächensprache(n)	Deutsch, Englisch, Schwedisch
PlugIns/ Anpassungsmög- lichkeiten	Bisher nicht vorhanden
Sonstiges	Ist noch im Beta Status, wird parallel im Rahmen eines ande- ren Init Projektes evaluiert

2.6 Erklärung der Tabellenkategorien

Anbei sind die einzelnen Dimensionen, die in den vorhergehenden Tabellen pro Werkzeug bewertet wurden, nochmals erklärt und beschrieben.

Anbieter	Name des Unternehmens/ Organisation, welche das Tool anbietet.
Land	Land in dem sich das Unternehmens/ Organisation befindet, welches das Tool anbietet





Anbieterwebsite	Die Internetpräsenz des Unternehmens/ Organisation, welches das Tool anbietet
Kosten	Die finanziellen Mittel, die für die Nutzung des Tools anfallen
Kollaborationsmöglichkeit	Die Möglichkeiten des Tools, ob und in welcher Art gemeinsam mit anderen Beteiligten an einem Projekt gearbeitet werden kann.
Fähigkeitslevel nach Fä- higkeitsskala	Das vom Tool geforderte Kompetenzlevel, um darin arbeiten zu können:
	1: Informiert - kann mitreden (Level Fahrstuhlgespräch)
	2: Qualifiziert - Zertifikat oder erste Erfahrung damit
	3: Praxiserfahren - das Wissen im Thema reicht aus, um es selbstständig anzuwenden
	4: Gereift - ausgeprägtes und tiefes Fachverständnis, hohe Expertise
Liste editierbarer Ontolo- giesprachen	Eine Auflistung der Ontologiesprachen die im Tool bearbeitet werden können
Bearbeitungsmöglichkeit	Hier werden die Möglichkeiten beschrieben, mit denen die Ontologien im Tool bearbeitet werden können. Dabei wird differenziert zwischen:
	 Textbasierter Bearbeitung ("Code")
	 Grafischer Bearbeitung auf Ebene einer visuellen Dar- stellung des Graphen
	3. "What you see, is what you get" Bearbeitung
SHACL möglich	Besteht die grundsätzliche Möglichkeit, SHACL im Tool anzuwenden
Unterstützung beim	Kann das Tool die Nutzenden beim Schreiben von SHACL
Schreiben von SHACL	unterstützen, z.B. durch Korrekturvorschläge, Syntax-Highlighting oder auto-completion.





Unterstützung beim De- buggen von SHACL	Kann das Tool die Nutzenden beim Debuggen von SHACL unterstützen, z.B. durch Hinweise auf Fehler
SPARQL möglich	Besteht die Möglichkeit, SPARQL Abfragen im Tool durchzuführen
Unterstützung beim Schreiben von SPARQL- Abfragen	Kann das Tool die Nutzenden beim Schreiben von SPARQL unterstützen, z.B. durch Korrekturvorschläge
Unterstützung beim Ausführen von SPARQL-Abfragen	Kann das Tool die Nutzenden beim Ausführen von SPARQL unterstützen
Unterstützung beim De- buggen von SPARQL-Ab- fragen	Kann das Tool die Nutzenden beim Debuggen von SPARQL unterstützen, z.B. durch eine Unterstützung einer "Schritt zu Schritt Abarbeitung" mit Haltepunkten
Offizielle Dokumentation vorhanden	Existiert eine offizielle Dokumentation für das Tool
Community-Support	Besteht eine aktive und entsprechend große Community für das Tool, die bei Bedarf Fragen zum Tool beantworten kann
Ist die Software frei ver- fügbar?	Wird das Tool unter einem Open Source Software Lizenz- modell oder einem anderem angeboten
Lizenz	Unter welcher genauen Lizenz das Tool angeboten wird
Import Möglichkeiten	Welche Formate kann das Tool importieren z.B. RDF, OWL etc.
Export Möglichkeiten	Welche Formate kann das Tool exportieren z.B. RDF, OWL etc.
Vorab beinhaltete Taxo- nomien/ Ontologien	Eine Auflistung der bereits vorab im Tool hinterlegten Taxo- nomien und Ontologien
Editierbarer Namespace: IRI/ URN etc.	Können die Namespaces, also die Identifikationsmerkmale z.B. IRI und URN im Tool angepasst werden





4-Augen Prinzip o.Ä. mög- lich	Besteht die Möglichkeit eingegebene Änderungen durch ein 4-Augen-Prinzip oder Vergleichbares abzusichern
Installationsort: Web oder Lokal?	Wird das Tool lokal auf dem Rechner der Nutzenden installiert oder ist es eine CLoudanwendung
Oberflächensprache(n)	In welchen Sprachen die für die Nutzenden sichtbaren Ele- mente des Tools verfügbar sind
PlugIns/ Anpassungsmög- lichkeiten	Ob PlugIns und vergleichbare Anpassungsmöglichkeiten bestehen
Sonstiges	Hier können Erkenntnisse aufgelistet werden, die keine der anderen Kategorien zugeordnet werden können.





3 Visualisierungstools

3.1 Einleitung

Visualisierungstools sind spezialisierte Softwareanwendungen, die darauf abzielen, komplexe ontologische Strukturen grafisch darzustellen. Diese Werkzeuge bieten eine visuelle Repräsentation der Beziehungen und Hierarchien innerhalb eines Wissensgraphen die ansonsten in formalen, oft für die normalen Nutzenden ohne tiefergehende Vorerfahrungen nur schwer verständlichen, maschinenlesbaren Formaten wie OWL oder RDF kodiert sind. Durch die visuelle Aufbereitung wird das Verständnis und die Analyse eines Wissensgraphen erheblich erleichtert, da Nutzende die Struktur und die Zusammenhänge innerhalb der Daten intuitiv erfassen können.

In dieser Übersicht werden verschiedene Visualisierungstools vorgestellt. Die Auswahl der Tools basiert auf ihrer Verbreitung, den angebotenen Funktionen und ihrer Benutzerfreundlichkeit. Ziel dieser Übersicht ist es, einen ersten Überblick über die verfügbaren Werkzeuge zu bieten und ihre Eignung für unterschiedliche Anwendungsfälle zu bewerten. Dabei werden sowohl etablierte Tools als auch neuere Entwicklungen betrachtet, um eine breite Palette von Optionen abzudecken.

3.2 ShowVoc

ShowVoc ist die Visualisierungskomponente von VocBench und somit ebenso wie VocBench eine Lösung des Publication Office der Europäischen Union. Sie dient dazu ontologische Strukturen grafisch darzustellen. Während ShowVoc grundsätzliche Funktionen für die Visualisierung bietet, zeigt sich das Tool in der Praxis als wenig ausgereift. Die Benutzeroberfläche wirkt veraltet und entspricht nicht den modernen Standards für Benutzerfreundlichkeit und Übersichtlichkeit. Dies kann die Nutzererfahrung erheblich beeinträchtigen, insbesondere bei intensiver Nutzung.

Ein bedeutendes Manko ist die eingeschränkte Anpassungsfähigkeit der Visualisierung. Beispielsweise lassen sich Farben und andere visuelle Elemente nur begrenzt anpassen, was die Flexibilität in der Darstellung von Wissensgraphen einschränkt. Diese Einschränkungen können insbesondere bei der Arbeit mit komplexen Datensätzen problematisch sein, da sie die Klarheit und Verständlichkeit der visuellen Darstellung beeinträchtigen.

Insgesamt bietet ShowVoc grundlegende Funktionen zur Visualisierung von Ontologien, jedoch sind die vorhandenen optischen und Benutzungs-Defizite sowie die begrenzten Anpassungsmöglichkeiten wesentliche Schwächen.





Tabellarische Bewertung ShowVoc

Anbieter	Publications Office der Europäischen Union
Land	EU
Anbieterwebsite	https://showvoc.uniroma2.it/doc/
Kosten	Keine
Kollaborationsmöglichkeit	Keine
Fähigkeitslevel nach Fä- higkeitsskala	3-4
Liste editierbarer Ontolo- giesprachen	RDF, OWL, SKOS, OntoLex, EDOAL
Bearbeitungsmöglichkeit	Textbasierter Bearbeitung ("Code")
Optische Anpassungs- möglichkeiten (Farbe, Form etc.)	Nein
Veränderbare Position der visuellen Elemente	Ja
Zoom Funktion	Ja
Erklärende Informations- felder möglich	Nein
Darstellung von SHACL möglich	Nein
Suchfunktion	Ja
Offizielle Dokumentation vorhanden	Ja <u>Digital Europa Thesaurus</u> , <u>Europass tables</u> , <u>ESCO taxonomies</u>





Community-Support	Es gibt eine <u>VocBench Nutzenden Gruppe</u> auf Google Groups. Nicht alle Fragen werden beantwortet, evtl. auf- grund der kleinen Gruppe an Nutzenden. Diese Gruppe kann auch für Anfragen in Bezug auf ShowVoc genutzt werden.
Ist die Software frei ver- fügbar?	Ja
Lizenz	Open Source Lizenz
Import Möglichkeiten	Keine genauen Formate genannt, es kann jedoch an Tripelstores angebunden werden, sodass gängige RDF Formate erwartbar sind. Aufgrund der Anknüpfungsmöglichkeit an VocBench sind die gleichen Formate zu erwarten
Export Möglichkeiten	SVG Datei
Vorab beinhaltete Taxo- nomien/ Ontologien	Bereits in ShowVoc bestehende EU Vokabulare und Ontologien wie Digital Europa Thesaurus, Europass tables, ESCO taxonomies können bei Bedarf genutzt werden
Editierbare Bezeichnun- gen	Nur auf Modellebene, also durch eine Anpassung des Codes, nicht auf der rein visuellen Ebene
4-Augen Prinzip o.Ä. mög- lich	Unklar
Installationsort: Web oder Lokal?	Lokal und Web möglich
Oberflächensprache(n)	Deutsch, Englisch, Spanisch, Französisch, Italienisch
PlugIns/ Anpassungsmög- lichkeiten	Keine
Sonstiges	SPARQL Abfragen sind möglich. Identische visuelle Darstellung wie in VocBench, jedoch mit einzelnen fehlenden Optionen da z.B. keine Möglichkeit be- steht einen Ausschnitt des Modells zu fixieren.





3.3 Stardog Explorer

Stardog Explorer gehört zu den aktuell am häufigsten eingesetzten Tools für die Visualisierung von Ontologien und semantischen Daten. Die Nutzergruppe umfasst zum Beispiel auch die OpenDVA Forschungsgruppe der Universität Jena. Stardog Explorer zeichnet sich durch umfangreiche visuelle Anpassungsmöglichkeiten und eine benutzerfreundliche Oberfläche aus, die es den Anwendern ermöglicht, komplexe Datenstrukturen individuell darzustellen.

Dies waren auch die Gründe, warum Stardog Explorer im Rahmen des RegCheck Projektes für Visualisierungen verwendet wurde.

Die Stärken von Stardog Explorer liegen in seiner Flexibilität bei der visuellen Darstellung und der modernen Optik.

Jedoch zeigt sich Stardog Explorer als geschlossenes System, in Kombination mit anderen Stardog Elementen, was die Integration und den Austausch von Daten nach außen erschwert. Die Exportmöglichkeiten sind stark begrenzt, sodass Visualisierungen in erster Linie innerhalb der Plattform verwendet werden können und nicht leicht in andere Formate oder Systeme überführt werden können. Diese Einschränkung bedeutet, dass das Tool weniger flexibel ist, wenn es um die Zusammenarbeit oder die Einbindung in heterogene Datenumgebungen geht. Auch ist wider Erwarten für eine kommerzielle Lösung der Einstieg in das Tool nicht intuitiv möglich. Die Dokumentation ist äußerst umfangreich, jedoch ist sie in einigen Bereichen so allgemein gehalten, dass sie nur begrenzt, hilfreich ist. Dies kann als ein Anreiz gesehen werden, auf kostenpflichtige Versionen zu wechseln, welche Supportservices beinhalten.

Tabellarische Bewertung Stardog Explorer

Anbieter	Stardog Union
Land	USA
Anbieterwebsite	https://www.stardog.com/
Kosten	Kostenlose Grundversion, mit verschiedenen Payversionen möglich
Kollaborationsmöglichkeit	Möglich, aber zugleich muss dafür auch der Zugriff auf die dahinterliegende Datenebene, die vorher in Stardog





eingespeist wurde, vergeben werden. Es müssen jedoch keine Bearbeitungsrechte an der Datenebene vergeben werden.

Fähigkeitslevel nach Fä- higkeitsskala	2
Liste editierbarer Ontolo- giesprachen	
Bearbeitungsmöglichkeit	Textbasiert und grafisch
Optische Anpassungs- möglichkeiten (Farbe, Form etc.)	Nein (nur in Stardog Designer)
Veränderbare Position der visuellen Elemente	Ja
Zoom Funktion	Ja
Erklärende Informations- felder möglich	Ja
Darstellung von SHACL möglich	Nein
Suchfunktion	Ja
Offizielle Dokumentation vorhanden	Ja, <u>Stardog Dokumentation</u>
Community-Support	Ja, aber eine kleine <u>Community</u> mit vergleichsweise wenig Aktivität
Ist die Software frei ver- fügbar?	Ja in der Basis-Variante, aber nicht Open Source
Lizenz	Software as a service





Import Möglichkeiten	N-Triples, RDF/XML, Turtle, TriG, N-Quads, JSON-LD (Jedoch muss jedes importierte Modell immer dem Stardog Schema angepasst werden)
Export Möglichkeiten	Tabellarischer Output der Daten
Vorab beinhaltete Taxo- nomien/ Ontologien	Keine
Editierbare Bezeichnun- gen	Ja
4-Augen Prinzip o.Ä. mög- lich	Nein
Installationsort: Web oder Lokal?	Web
Oberflächensprache(n)	Englisch
PlugIns/ Anpassungsmög- lichkeiten	Schnittstellen zu einer Vielzahl von Datenbanken möglich
Sonstiges	Nutzende müssen auf gleichen Datenbasis in Stardog Studio zugreifen können, um etwa die Visualisierung im Tool sehen zu können

3.4 Sparnatural

Sparnatural ist primär als Benutzeroberfläche für SPARQL Anfragen konzipiert. Das Tool bietet eine benutzerfreundliche Schnittstelle, die es ermöglicht, Abfragen an Wissensgraphen zu stellen, ohne dass die Anwender selbst SPARQL-Code schreiben müssen. Diese Funktionalität ist besonders vorteilhaft für Personen, die mit semantischen Datenmodellen und Wissensgraphen arbeiten möchten, jedoch nicht mit den technischen Aspekten der Abfragesprache vertraut sind.





Es ist dabei wichtig zu betonen, dass Sparnatural nicht als klassisches Visualisierungstool für Wissensgraphen selbst konzipiert ist. Das Tool bietet keine Funktionen zur grafischen Darstellung der Datenstrukturen innerhalb eines Wissensgraphen. Stattdessen konzentriert es sich auf die Zugänglichkeit und Benutzerfreundlichkeit der Datenabfrage. Somit visualisiert bzw. erleichtert es in erster Linie die Abfrage an den Wissensgraphen, weniger den Wissensgraphen selbst.

Trotz der fehlenden Visualisierungsfunktionen stellt Sparnatural ein wertvolles Tool dar, insbesondere in Szenarien, in denen der Fokus auf der Datenabfrage und -analyse liegt. Die intuitive Benutzeroberfläche trägt zur Vereinfachung des Zugriffs auf semantische Daten bei und erleichtert die Interaktion mit Wissensgraphen, wodurch sie für eine breitere Nutzerbasis zugänglich wird.

Tabellarische Bewertung Sparnatural

Anbieter	SPARNA
Land	Frankreich
Anbieterwebsite	https://sparnatural.eu/
Kosten	Kostenlos
Kollaborationsmöglichkeit	Keine
Fähigkeitslevel nach Fä- higkeitsskala	1
Liste editierbarer Ontolo- giesprachen	Keine
Schnittstellen zu Edi- tiertools	Keine
Bearbeitungsmöglichkeit	"What you see, is what you get" Bearbeitung
Optische Anpassungs- möglichkeiten (Farbe, Form etc.)	Farben können angepasst werden





Veränderbare Position der visuellen Elemente	Nein
Zoom Funktion	Nein
Erklärende Informations- felder möglich	Ja
Darstellung von SHACL möglich	Nein
Suchfunktion	Ja
Offizielle Dokumentation	Ja
vorhanden	https://docs.sparnatural.eu/
Community-Support	Nein
Ist die Software frei ver- fügbar?	Ja
Lizenz	LGPL-3.0
Import Möglichkeiten	Es kann an bestehende SPARQL Endpoints angeknüpft werden und ist unabhängig vom Datenformat
Evnort Möglichkoiton	
Export Möglichkeiten	Keine
Vorab beinhaltete Taxo- nomien/ Ontologien	Keine
Vorab beinhaltete Taxo-	
Vorab beinhaltete Taxo- nomien/ Ontologien Editierbare Bezeichnun-	Keine





Oberflächensprache(n)	Englisch
lichkeiten	Keine
Sonstiges	

3.5 OWLGrEd

OWLGrEd, entwickelt von der Universität Litauen, war insbesondere in akademischen Kreisen ein zeitweise populäres Tool und wurde in zahlreichen Forschungsarbeiten erwähnt.

Zum jetzigen Zeitpunkt ist OWLGrEd jedoch veraltet und entspricht nicht mehr den aktuellen Standards der Benutzerfreundlichkeit. Die Benutzeroberfläche ist antiquiert und wenig intuitiv, was die Handhabung des Tools erheblich erschwert. Zudem sind die Anpassungsmöglichkeiten stark eingeschränkt, was die Flexibilität bei der Erstellung und Anpassung von Visualisierungen begrenzt. Ein weiteres bedeutendes Manko ist der begrenzte Support, der nur noch in sehr eingeschränktem Umfang verfügbar ist. Die Kombination aus veralteter Technologie, eingeschränkter Anpassungsfähigkeit und mangelndem Support macht OWLGrEd zunehmend ungeeignet für aktuelle Anforderungen an ein Visualisierungstools.

Tabellarische Bewertung OWLGrEd

Anbieter	Institute of Mathematics and Computer Science, Universität Litauen
Land	Litauen
Anbieterwebsite	http://owlgred.lumii.lv/
Kosten	Kostenlos
Kollaborationsmöglichkeit	Nein





Fähigkeitslevel nach Fä- higkeitsskala	2
Liste editierbarer Ontolo- giesprachen	OWL, RDF
Schnittstellen zu Edi- tiertools	Nein
Bearbeitungsmöglichkeit	"What you see, is what you get" Bearbeitung
Optische Anpassungs- möglichkeiten (Farbe, Form etc.)	Formen und Farben sind anpassbar, auch eigene Icons sind möglich
Veränderbare Position der visuellen Elemente	Ja
Zoom Funktion	Ja
Erklärende Informations- felder möglich	Ja, in Form von Kommentaren in der Visualisierung
Darstellung von SHACL möglich	Nein
	Nein Ja, aber nicht funktional
möglich	
möglich Suchfunktion Offizielle Dokumentation	Ja, aber nicht funktional Ja, jedoch nur sehr minimalistisch
möglich Suchfunktion Offizielle Dokumentation vorhanden	Ja, aber nicht funktional Ja, jedoch nur sehr minimalistisch http://owlgred.lumii.lv/get_started
möglich Suchfunktion Offizielle Dokumentation vorhanden Community-Support Ist die Software frei ver-	Ja, aber nicht funktional Ja, jedoch nur sehr minimalistisch http://owlgred.lumii.lv/get_started Nein





Export Möglichkeiten	SVG
Vorab beinhaltete Taxo- nomien/ Ontologien	Keine
Editierbare Bezeichnun- gen	Nein
4-Augen Prinzip o.Ä. mög- lich	Nein
Installationsort: Web oder Lokal?	Lokal
Oberflächensprache(n)	Englisch
PlugIns/ Anpassungsmög- lichkeiten	Möglich, eigene Bibliothek ist vorhanden http://owlgred.lumii.lv/extensions
Sonstiges	In den vergangenen Jahren nicht mehr geupdatet, sehr veraltete und unattraktive Benutzeroberfläche

3.6 WebVOWL

WebVOWL ist ein zeitweise häufig verwendetes Visualisierungstool innerhalb akademischer Publikationen gewesen. Das Tool ermöglichte eine grundlegende Visualisierung von Tripeln.

Jedoch wurden seit 2022 keine Updates mehr für WebVOWL veröffentlicht, was Bedenken hinsichtlich der langfristigen Wartung und Aktualität des Tools aufwirft. Auch die dahinterliegende Organisation, Visual Data Web, hat ihren Betrieb eingestellt.

Zudem bietet WebVOWL nur begrenzte Möglichkeiten zur Anpassung der Visualisierungen innerhalb der Plattform. Dies bedeutet, dass die Flexibilität in der Darstellung der Tripel und deren Beziehungen stark eingeschränkt sind, was die Anpassung der Visualisierungen an spezifische Anforderungen oder Präferenzen erschwert.





WebVOWL ist somit zwar noch thematisch immer wieder präsent, jedoch für eine zukünftige Nutzung aktuell nicht empfehlenswert, sofern es nicht von einer neuen Organisation wieder belebt wird.

Tabellarische Bewertung WebVOWL

Anbieter	Visual Data Web (nicht mehr aktiv)
Land	USA
Anbieterwebsite	https://github.com/VisualDataWeb/WebVOWL
Kosten	Keine
Kollaborationsmöglichkeit	Keine
Fähigkeitslevel nach Fä- higkeitsskala	2
Liste editierbarer Ontolo- giesprachen	Unklar
Schnittstellen zu Edi- tiertools	Nein
Bearbeitungsmöglichkeit	Bisher nur in einer experimentellen Version möglich
Optische Anpassungs- möglichkeiten (Farbe, Form etc.)	Nein
Veränderbare Position der visuellen Elemente	Ja
Zoom Funktion	Ja
Erklärende Informations- felder möglich	Nein





Darstellung von SHACL möglich	Nein
Suchfunktion	Ja
Offizielle Dokumentation vorhanden	Nein
Community-Support	Nein
Ist die Software frei ver- fügbar?	Ja
Lizenz	MIT License (MIT)
Import Möglichkeiten	URI, ansonsten unklar
Export Möglichkeiten	JSON, SVG, TeX, TTL, URL
Vorab beinhaltete Taxo- nomien/ Ontologien	Keine
Editierbare Bezeichnun- gen	Nein
4-Augen Prinzip o.Ä. mög- lich	Nein
Installationsort: Web oder Lokal?	Lokal
Oberflächensprache(n)	Englisch
PlugIns/ Anpassungsmög- lichkeiten	Nein
Sonstiges	





3.7 Erklärung der Tabellenkategorien

Anbei sind die einzelnen Dimensionen, die in den vorhergehenden Tabellen pro Werkzeug bewertet wurden, nochmals erklärt und beschrieben.

Anbieter	Name des Unternehmens/ Organisation, welche das Tool anbietet.
Land	Land in dem sich das Unternehmens/ Organisation befindet, welches das Tool anbietet
Anbieterwebsite	Die Internetpräsenz des Unternehmens/ Organisation, welches das Tool anbietet
Kosten	Die finanziellen Mittel, die für die Nutzung des Tools anfallen
Kollaborationsmöglichkeit	Die Möglichkeiten des Tools, ob und in welcher Art gemeinsam mit anderen Beteiligten an einem Projekt gearbeitet werden kann.
Fähigkeitslevel nach Fä- higkeitsskala	Das vom Tool geforderte Kompetenzlevel, um darin arbeiten zu können:
	1: Informiert - kann mitreden (Level Fahrstuhlgespräch)
	2: Qualifiziert - Zertifikat oder erste Erfahrung damit
	3: Praxiserfahren - das Wissen im Thema reicht aus, um es selbstständig anzuwenden
	4: Gereift - ausgeprägtes und tiefes Fachverständnis, hohe Expertise
Liste editierbarer Ontolo- giesprachen	Eine Auflistung der Ontologiesprachen die im Tool bearbei- tet werden können
Bearbeitungsmöglichkeit	Hier werden die Möglichkeiten beschrieben, mit denen die Visualisierungen bearbeitet werden können. Dabei wird dif- ferenziert zwischen:
	Textbasierter Bearbeitung ("Code")
	2. Grafischer Bearbeitung auf Ebene einer visuellen Dar-

stellung des Graphen

33









"What you see, is what you get" Bearbeitung

Optische Anpassungs- möglichkeiten (Farbe, Form etc.)	Inwiefern können innerhalb der Darstellung des Wissensgra- phen Farben, Formen etc. der Nodes und Kanten optisch verändert, gruppiert etc. werden.
Veränderbare Position der visuellen Elemente	Können die einzelnen Objekte bewegt und ihre Position zu- einander verändert werden
Zoom Funktion	Ob die Ansicht vergrößert oder verkleinert werden kann.
Erklärende Informations- felder möglich	In welchem Umfang können innerhalb der Visualisierung er- klärende Informationsfelder eingeblendet werden
Darstellung von SHACL möglich	Lassen sich SHACL Beschränkungen visuell darstellen
Suchfunktion	Kann innerhalb der Visualisierung nach Nodes o.Ä. gesucht werden
Offizielle Dokumentation vorhanden	Existiert eine offizielle Dokumentation für das Tool Hier soll auch die Dokumentation verlinkt sein
Community-Support	Besteht eine aktive und entsprechend große Community für das Tool, die bei Bedarf Fragen zum Tool beantworten kann
Ist die Software frei ver- fügbar?	Wird das Tool unter einem FOSS (Free Open Source Software) Lizenzmodell oder einem anderem angeboten
Lizenz	Unter welcher genauen Lizenz das Tool angeboten wird
Import Möglichkeiten	Welche Formate kann das Tool importieren z.B. RDF, OWL etc.
Export Möglichkeiten	Welche Formate kann das Tool exportieren z.B. PDF, Powerpoint, PNG, JPEG etc.
Vorab beinhaltete Taxo- nomien/ Ontologien	Eine Auflistung der bereits vorab im Tool hinterlegten Taxo- nomien und Ontologien





Editierbare Bezeichnun- gen	Können die Namen in der Visualisierung, abweichend zur eigentlichen Benennung in der Ontologie, verändert werden. um z.B. für Fachfremde leichter verständlich zu sein
4-Augen Prinzip o.Ä. mög- lich	Besteht die Möglichkeit eingegebene Änderungen durch ein 4-Augen-Prinzip oder Vergleichbares abzusichern
Installationsort: Web oder Lokal?	Wird das Tool lokal auf dem Rechner der Nutzenden installiert oder ist es eine Cloudanwendung
Oberflächensprache(n)	In welchen Sprachen die für die Nutzenden sichtbaren Ele- mente des Tools verfügbar sind
Plugins/ Anpassungsmög- lichkeiten	Ob PlugIns und vergleichbare Anpassungsmöglichkeiten bestehen.
Sonstiges	Hier können Erkenntnisse aufgelistet werden, die keine der anderen Kategorien zugeordnet werden können.
Anbieter	Name des Unternehmens/ Organisation, welche das Tool anbietet.
Land	Land in dem sich das Unternehmens/ Organisation befindet, welches das Tool anbietet





4 Tripelstores

4.1 Einleitung

Tripelstores sind spezialisierte Datenbanken, die entwickelt wurden, um semantische Daten effizient zu speichern, verwalten und abzufragen. Die dort hinterlegten Daten werden in Form von Tripeln gespeichert, die Aussagen über Ressourcen in einem maschinenlesbaren Format darstellen. Tripelstores sind ein integraler Bestandteil von Anwendungen, die auf dem Semantic Web basieren, um die oft großen Mengen an Tripeln zu hinterlegen.

Tripelstores sind so konzipiert, dass sie große Mengen semantischer Daten verarbeiten können. Im Gegensatz zu relationalen Datenbanken, die Tabellen und Zeilen verwenden, um Daten zu speichern, organisieren Tripelstores ihre Daten in einem Netzwerk von Knoten und Kanten, also einer Graphstruktur. Diese Graphstruktur ermöglicht es, komplexe Beziehungen zwischen Datenpunkten zu modellieren und zu analysieren.

Im Gegensatz zu tabellarischen Datenbanken sind Tripelstores in der Lage SPARQL-Anfragen zu verarbeiten. SPARQL ist eine Abfragesprache, die speziell für die Navigation und das Abfragen von RDF-Daten (Ressource Description Framework, einem auf Tripeln basierenden Standard des World Wide Web Consortiums) entwickelt wurde. Sie ermöglicht es, gezielte Anfragen zu stellen, die komplexe Beziehungen zwischen den Daten aufdecken können.

Da Tripelstores auf die Abfrage von Tripeln optimiert sind, erzielen sie insbesondere bei der Verarbeitung von Abfragen, die große und komplexe Tripel-Datensätze betreffen sehr gute Ergebnisse in Bezug auf die Suchgeschwindigkeit und die Effizienz bei der Abfrageausführung.

Viele Tripelstores implementieren auch Mechanismen zur Datenkompression und zur Verwaltung von in Tripeln gespeicherten Daten, um die Skalierbarkeit zu erhöhen.

Graph Datenbanken sind Tripelstores sehr ähnlich. Auch sie speichern Daten in einer Graphform, jedoch nicht spezifisch basierend auf Tripeln. Daher sind sie jedoch nicht im gleichen Maße auf die Verarbeitung von Tripeln spezialisiert, sodass sie im Durchschnitt eine geringere Effizienz in Bezug auf die Speicherung von Tripeln aufweisen. Auch ergänzende Funktionalitäten wie die Unterstützung von komplexen Inferenzen sind weniger oder gar nicht vorhanden. Daher wurde im Falle dieser Analyse der Fokus auf Tripelstores gelegt, ohne jedoch grundsätzlich Graph Datenbanken außen vor zu lassen.

Im Rahmen unserer Recherche zu Triple Stores, die für die Lagerung von Registermetadaten in Wissensgraphen in Frage kommen, hat sich gezeigt, dass die relevanten Unterschiede zwischen verschiedenen Triple Store-Systemen erst bei sehr großen Datenmengen signifikant werden. Insbesondere bei der Verarbeitung von mehreren Millionen Tripeln treten Differenzen in der Performance, Zuverlässigkeit und Skalierbarkeit deutlicher zutage.

Dabei ist zu beachten, dass verschiedene Triple Stores in unterschiedlichen Bereichen ihre spezifischen Stärken und Schwächen haben. Beispielsweise kann ein System sich durch eine





höhere Abfragegeschwindigkeit auszeichnen, während ein anderes System möglicherweise weniger fehlerhafte oder abgebrochene Suchanfragen generiert. Diese Unterschiede sind jedoch kontextabhängig und variieren je nach Anwendungsfall und den spezifischen Anforderungen des jeweiligen Einsatzbereichs.¹

Für die Anwendungsszenarien von RegCheck sind diese Unterschiede derzeit von untergeordneter Bedeutung. Die Datenmengen, die wir im Rahmen dieses Projekts erwarten, liegen deutlich unter der Schwelle, bei der diese Unterschiede eine entscheidende Rolle spielen würden. Auch im Falle einer zukünftigen Implementation des Wissensgraphen in der Fläche, bei der eine deutlich größere Anzahl von Registermetadaten integriert werden könnte, ist es fraglich, ob ein Bereich erreicht wird, in dem die Leistungsunterschiede zwischen den Triple Stores ausschlaggebend wären.

Ein weiteres Hindernis bei der Entscheidungsfindung und des Vergleichs von Tripelstores ist der aktuelle Stand der Forschungsliteratur zu Triple Stores. Es hat sich gezeigt, dass die vorhandene Literatur nicht mehr den neuesten technischen Entwicklungen entspricht. Das Thema Triple Stores hat in der Praxis sowie in der akademischen Forschung zusammen mit Ontologien und Wissensgraphen in jüngster Zeit wieder an Bedeutung gewonnen, jedoch gibt es derzeit nur wenig aktuelle, unabhängige Forschung, die den technischen Fortschritt und die Leistungsfähigkeit moderner Triple Stores umfassend untersucht. Die vorhandenen Informationen werden häufig von den Anbietern der jeweiligen Systeme beeinflusst, was eine objektive Bewertung erschwert.

Vor diesem Hintergrund ist es entscheidend, einen Überblick der relevantesten Triple Stores zu gewinnen. Daher stellen wir nachfolgend einige der bekanntesten und leistungsfähigsten Triple Stores vor, die gegenwärtig auf dem Markt verfügbar sind. Hier werden die spezifischen Rahmenbedingungen sowie die jeweiligen Vor- und Nachteile vorgestellt.

Im Anschluss an dieses Forschungsprojekt, sobald ersichtlich ist, welche Registermetadaten und welcher Umfang an Tripeln in einem nächsten Schritt angebunden werden, ist eine vertiefende anbieterunabhängige technische Vertestung der einzelnen Triple Stores empfehlenswert. Eine solche Evaluierung kann dann durch gezielte Leistungstests und Vergleichsstudien auf Basis der vorliegenden Anforderungen erfolgen, um die optimale Lösung für die Anwendung des Wissensgraphen in der öffentlichen Verwaltung zu identifizieren.

4.2 Stardog

Stardog ist ein Triplestore, der von der Stardog Union, einem Softwareunternehmen mit Sitz in den USA, entwickelt und vertrieben wird. Das Unternehmen wurde 2012 gegründet und hat sich auf die Entwicklung von Wissensgraphen-Technologien spezialisiert, die es Unternehmen

¹ Vergleiche hierzu zum Beispiel: Atemezing, G.A., Amardeilh, F. (2018). Benchmarking Commercial RDF Stores with Publications Office Dataset, abrufbar unter https://doi.org/10.1007/978-3-319-98192-5 54 (Zuletzt abgerufen am 23.08.2024)





ermöglichen, ihre Datenbestände semantisch zu verknüpfen und zu analysieren. Stardog ist besonders bekannt für seine Fähigkeiten zur Datenintegration und -visualisierung sowie für seine Unterstützung von Inferenz und Ontologiemanagement.

Stardog ist eine kommerzielle Softwarelösung, die in verschiedenen Lizenzmodellen angeboten wird, um den unterschiedlichen Bedürfnissen von Unternehmen und Organisationen gerecht zu werden. Im Rahmen des RegCheck Projektes wurde die Stardog Free Version verwendet, die die grundlegenden Funktionalitäten für kleinere Projekte und Evaluierungen bietet. Diese Version ist jedoch in Bezug auf die Größe der verwaltbaren Daten eingeschränkt. Aber auch hier ist die maximale Begrenzung von Tripeln mit ca. einer Million vollkommen ausreichend. Es unterstützt RDF und SPARQL vollständig und erweitert diese Standards durch leistungsfähige Inferenz, Ontologiemanagement und Datenvisualisierung.

Stardog bietet umfassende Unterstützung für SPARQL 1.1, einschließlich Updates, Inferenz und Transaktionen. Darüber hinaus bietet Stardog in seinen mit dem Tripelstore verbundenen Visualisierungs- und Bearbeitungsoberflächen die Möglichkeit SPARQL Abfragen mithilfe eines nutzerfreundlicheren Interfaces durchzuführen.

Auch bietet Stardog in der Theorie umfassende Integrationsmöglichkeiten und Schnittstellen. Zugleich ist es jedoch der Ansatz von Stardog, dass es als eine komplette Lösung gedacht ist, welche sowohl Datenlagerung, Visualisierung als auch Auswertung in einem geschlossenen System vorsieht. Die Nutzung als reiner Tripelstore ist nicht Hauptfunktionalität von Stardog, sondern der Tripelstore ist Teil eines Verbundes der Stardog-Funktionalitäten. Ebenso ist das Teilen von Daten mit Personen, die nicht Stardog Nutzende sind, aktuell nicht vorgesehen. Somit ist Stardog zwar eine an Funktionalitäten starke Lösung, birgt jedoch die Gefahr eines Vendor-lock in.

4.3 Graph DB

GraphDB ist eine Graph Datenbank, entwickelt von Ontotext, einem führenden Anbieter semantischer Technologien mit Sitz in Bulgarien. Ontotext ist aktuell stark präsent auf wissenschaftlichen Konferenzen und Veranstaltungen, wie beispielsweise dem Leipziger Semantic Web Day, und positioniert sich als aktiver Akteur im Bereich der semantischen Technologien in Europa. Ontotext als Unternehmen selbst ist bisher stärker auf Industrieanwendungen fokussiert.

Es werden verschiedene Editionen angeboten, die unterschiedliche Funktionalitäten und Preisstrukturen umfassen. Die GraphDB Free Version ist kostenlos verfügbar und bietet grundlegende Triple-Store-Funktionen, die für viele Anwendungsfälle ausreichend sind. Sie eignet sich besonders für kleinere Projekte, Forschungszwecke oder als Einstiegslösung. Erweiterte Versionen umfassen spezielle Supportpakete, erweitertes Caching etc. Auch verschiedene Cloud Optionalitäten sind in diesen kostenpflichtigen Versionen verfügbar-

GraphDB ist auf die effiziente Speicherung und Verwaltung von RDF-Daten optimiert. Es unterstützt die vollständige Implementierung des RDF 1.1-Standards und bietet umfassende





SPARQL 1.1-Unterstützung, was es zu einem hochgradig kompatiblen und leistungsfähigen Triple Store macht.

Diese Lösung lässt sich vergleichsweise einfach in bestehende IT-Infrastrukturen integrieren. Es bietet APIs für Java, REST, und SPARQL-Endpunkte, die es ermöglichen, GraphDB mit anderen Anwendungen und Systemen zu verbinden. Darüber hinaus unterstützt es die Integration mit Ontologie-Editoren, ETL-Tools und Business-Intelligence-Plattformen.

GraphDB bietet außerdem Plugins für erweiterte Funktionen, wie beispielsweise benutzerdefinierte Inferenzregeln, Textsuche und Verknüpfungen mit externen Datenquellen. Diese Modularität ermöglicht es Nutzenden, GraphDB an spezifische Anwendungsfälle anzupassen. Die verschiedenen Editionen ermöglichen es, GraphDB sowohl in kleinen Projekten als auch in größeren Kontexten anzuwenden bzw. bei Bedarf zu erweitern. Vor allem zeichnet sich Graph DB auch durch einen sehr benutzerfreundlichen Einstieg in die Nutzung aus. Dies gelingt durch detaillierte geführte Tutorials und Dokumentationen. Damit hebt sich GraphDB stark von anderen Lösungen in diesem Bereich ab, welche stärkere Vorkenntnisse erfordern.

4.4 RDF₄J

RDF4J wurde ursprünglich unter dem Namen OpenRDF Sesame von Aduna, einem niederländischen Unternehmen, entwickelt und später unter die Schirmherrschaft der Eclipse Foundation gestellt. Die Software ist eine Open-Source-Lösung und wird unter der Eclipse Public License (EPL) bereitgestellt. Diese Lizenz ermöglicht es, RDF4J kostenfrei zu nutzen, zu modifizieren und in eigenen Projekten einzusetzen. Die Open-Source-Natur von RDF4J hat zu einer breiten Akzeptanz in der akademischen Gemeinschaft.

Auch die Grundeinstellungen für die Installation von VocBench auf eigenen Servern gehen von einem RDF4J Tripelstore aus.

RDF4J ist kostenlos erhältlich, da es unter einer Open-Source-Lizenz veröffentlicht wird. Es fallen keine Lizenzgebühren an. Jedoch können zusätzliche Kosten für die Implementierung, Wartung und Schulung von Mitarbeitern entstehen, insbesondere wenn RDF4J in groß angelegten oder produktiven Umgebungen eingesetzt wird.

RDF4J kann sowohl als relationale Datenbank oder als Triple Store genutzt werden. Diese Flexibilität ermöglicht es, RDF4J sowohl für kleine, speicherinterne Anwendungen als auch für große, skalierbare Projekte einzusetzen.

RDF4J bietet vollständige Unterstützung für SPARQL. Dies umfasst komplexe Abfragen, Updates, Inferenz und bei Bedarf die Unterstützung für Named Graphs. Im Gegensatz zu anderen Tripelstores unterstützt RDF4J jedoch nur SPARQL und SeRQL. Dies stellt jedoch kein Problem dar, da im Rahmen von semantischen Anwendungen SPARQL die relevanteste Abfragesprache darstellt.

RDF4J kann vergleichsweise einfach in Java-basierte Anwendungen integriert werden, was durch seine APIs und modulare Architektur unterstützt wird. Die Software ermöglicht die





einfache Integration von benutzerdefinierten Speicherlösungen oder Inferenzmaschinen. Somit kann RDF4J mühelos mit anderen Tripelstores verbunden und kombiniert werden.

RDF4J stellt eine robuste und flexible Open Source Lösung als Triple Store dar, die sich vor allem durch ihre hohe Anpassungsfähigkeit auszeichnet. Dank der Open-Source-Lizenzierung entstehen keine direkten Kosten, was RDF4J attraktiv macht, jedoch auch die Möglichkeiten für Support einschränkt.

4.5 TypeDB

TypeDB, früher bekannt als Grakn, ist eine Wissensdatenbank, die von Vaticle entwickelt wird, einem Unternehmen mit Sitz in London, Großbritannien. Vaticle hat TypeDB als ein stark typisiertes Wissensgraph-Management-System entwickelt, das über herkömmliche Triple Stores hinausgeht, indem es eine andere und in Teilen komplexere Modellierung von Daten anbietet. TypeDB unterscheidet sich von traditionellen Triple Stores durch seine Fähigkeit, komplexe Beziehungen und Abfragen über ein erweitertes Konzept von Entitäten, Beziehungen und Regeln zu handhaben.

TypeDB ist als Open-Source-Software unter der Apache 2.0 Lizenz kostenlos verfügbar, was bedeutet, dass es ohne Lizenzkosten genutzt und modifiziert werden kann. Vaticle bietet jedoch auch eine kommerzielle Version an, die zusätzliche Features, Optimierungen und professionellen Support umfasst. Die genauen Kosten für die kommerzielle Version variieren je nach Umfang der Implementierung, der benötigten Unterstützung und der spezifischen Anwendungsanforderungen.

Obwohl TypeDB nicht als klassischer Triple Store betrachtet wird, erfüllt es viele ähnliche Funktionen und bietet erweiterte Möglichkeiten zur Modellierung und Abfrage von Daten. TypeDB nutzt ein stark typisiertes Datenmodell, das es ermöglicht, erweiterte Strukturen darzustellen, die über Subjekt-Prädikat-Objekt-Triples hinausgehen. Dies ist jedoch nicht relevant für die Anwendungsfälle von RegCheck im Meldewesen.

TypeDB unterstützt nicht SPARQL, sondern verwendet stattdessen eine eigene Absprachesprache: TypeQL. TypeQL ist speziell dafür konzipiert, Abfragen zu ermöglichen, die von Triple-Mustern abweichen bzw. einer anderen Logik folgen. TypeQL erlaubt es, strukturierte Abfragen zu formulieren, die auf dem eigenen semantischen Modell von TypeDB basieren.

Somit ist Type DB zum jetzigen Zeitpunkt keine präferierte Lösung im Kontext der Anwendung von semantischen Daten in der öffentlichen Verwaltung. Es muss jedoch beachtet werden, dass TypeDB neu ist (2024 erst in den stable release gekommen) und weitere Entwicklungen zu erwarten sind.





5 Fazit und Ausblick

Zusammenfassend kann vor dem Hintergrund der drei bewerteten Arten von Softwarekomponenten (Editiertools, Triplestores, Visualisierungstools) festgehalten werden:

Editiertools sollen zur Pflege des Wissensgraphen und zum Einspeisen neuer Daten verwendet werden. Im Vergleich der Editortools stehen wir insgesamt vor der Herausforderung von veralteten Benutzeroberflächen und nicht mehr gepflegten PlugIns, wie z.B. bei Protégé und OWLGrEd. Auch hohe Einstiegshürden wie im Falle von VocBench stellen ein Problem dar. Webprotégé ist dahingegen sowohl in Bezug auf Benutzung als auch Oberfläche eine positive Ausnahme. Es ist das aktuell bestnutzbare Tool, trotz der in Teilen eingeschränkten Funktionalitäten. Zugleich muss berücksichtigt werden, dass selbst solche, im Vergleich zur Bearbeitung im Quelltexteditor, hürdenärmeren Editiertools immer noch ein gewisses Grundverständnis der Mitarbeitenden, voraussetzt.

Nichtsdestotrotz sind die aufgezeigten Editoren leichter, verständlicher und für eine größere Benutzergruppe verwendbar als die im Rahmen von RegCheck verwendeten Quelltexteditoren, welche tiefergehendes Wissen sowohl über Ontologien als auch das Schreiben von Quellcode voraussetzen.

Im Falle von **Tripelstores**, werden, wie bereits beschrieben, die relevanten Unterschiede zwischen verschiedenen Triple Store-Systemen erst bei sehr großen Datenmengen signifikant. Es treten erst dann Differenzen in der Performance, Zuverlässigkeit und Skalierbarkeit deutlicher zutage, welche auch abhängig sind von den konkreten Nutzungsszenarien des Tripelstores. Wichtig ist daher eine Erprobung der Tripelstores abhängig vom späteren Umfang der Implementierung des RegCheck Modells.

Zum jetzigen Zeitpunkt kann daher kein absoluter Favorit herausgearbeitet werden. In einem späteren umfangreicheren Vergleich sollte Stardog, aufgrund seiner integrierten Funktionalitäten wie z.B. der dazugehörigen Visualisierungsmöglichkeiten, und GraphDB, aufgrund der niedrigen Einstiegshürden mit sehr nutzerfreundlichen Tutorials und praktischer Benutzeroberfläche, besonders berücksichtigt werden. Insbesondere GraphDB hat hierbei den zusätzlichen Vorteil, dass es ein europäischer Anbieter ist. Für den PoC RegCheck wurde der Triplestore auf Basis von Stardog umgesetzt.

In der Betrachtung der **Visualisierungstools** zeigte sich, dass für Nutzende der öffentlichen Verwaltung ein Tool nützlich wäre, das intuitive Abfragen ermöglicht, ohne tiefgehendes Wissen über Wissensgraphen vorauszusetzen.

Die Fähigkeit, Daten aus Wissensgraphen abzurufen und zu analysieren, sollte dabei möglichst benutzerfreundlich und zugänglich gestaltet werden, um den Bedürfnissen von Anwendern gerecht zu werden, die nicht mit den technischen Details der semantischen Datenverarbeitung vertraut sind. Ein geeignetes Tool sollte daher nicht nur die Abfragefunktionalitäten bereitstellen, sondern diese auch auf eine Weise präsentieren, die die Benutzererfahrung optimiert. Diese Funktionalitäten werden aktuell von bestehenden Visualisierungstools nur unzureichend erfüllt.









In der grundlegendsten Ausführung wird ein einfaches SPARQL Nutzer Interface benötigt, welche es den Nutzenden ermöglicht, ihre spezifischen Informationsbedürfnisse zu adressieren, ohne sich in die Komplexität der SPARQL-Syntax einarbeiten zu müssen. Zugleich sollte es jedoch weitergehend als eine reine SPARQL Abfrage sein, um ein "Stöbern" innerhalb des Wissensgraphen zu erlauben und Querverbindungen und neue Erkenntnisse zu ermöglichen.

Ein Werkzeug, welches einem Teil dieser Anforderungen gerecht wird, ist Sparnatural. Wie bereits dargestellt, bietet Sparnatural eine benutzerfreundliche Oberfläche für SPARQL-Abfragen. Sie ermöglicht, ohne tiefgehende Kenntnisse der Abfragesprache, mit dem Wissensgraphen zu interagieren. Für den PoC RegCheck wurde der die Visualisierung mittels Sparnatural umgesetzt. Obwohl Sparnatural über keine Visualisierungsfunktionen für den Wissensgraphen verfügt, stellt es eine Lösung für die Abfrage und die Analyse von Daten dar, indem es eine einfache und zugängliche Benutzeroberfläche bereitstellt. Zugleich ist jedoch die Möglichkeit des "Stöbern", also entlang von visualisierten Verbindungen neue Querverbindungen zu entdecken, nicht möglich. Als Ergebnis der Analysen und Test konnte aber auch festgestellt werden, dass keines der untersuchten und aktuell auf dem Markt verfügbaren Tools dies aktuell leisten kann.

Für zukünftige Projekte und Arbeiten ist eine tiefergehende Auseinandersetzung mit Lösungen, die im Rahmen des erstarkenden allgemeinen Interesses an Wissensgraphen neu entwickelt werden, notwendig. In diesem Zusammenhang sollten auch Eigenentwicklungen in Betracht gezogen werden. Die Schaffung maßgeschneiderter Lösungen, die die spezifischen Anforderungen und Bedürfnisse der Nutzenden in der öffentlichen Verwaltung adressieren, können dazu beitragen, die Interaktion mit semantischen Datenmodellen weiter zu verbessern und den Zugang zu komplexen Wissensgraphen zu vereinfachen. Hierdurch würde es möglich, an die verschiedenen Nutzergruppen angepassten Funktionalitäten zu entwickeln.

Zusammenfassend lässt sich festhalten, dass die Entwicklung von intuitiven Abfragetools für Wissensgraphen einer der nächsten wichtigen Schritte für die Anwendung und Implementation von semantischen Lösungen im Kontext der Registermodernisierung darstellt.