Semantic Web

Investitionen in grüne Energie der Bundesregierung Deutschland

Sebastian Pötter

Jahrgang: 20INM

im Studiengang
Informatik Master

Semantic Web
Prof. Thomas Riechert

25. August 2021

Inhaltsverzeichnis

1 Einleitung **1**

2 Datenquellen **1**
2.1 Verteilung der Erzeugung 1
2.2 Wahlen in Deutschland bis 2017 2
2.3 Investition in erneuerbaren Energie bis 2020 2
2.4 Investitionen in Erzeugungsanlagen 3

3 Arbeitsvorgang **3**
3.1 Datensuche 3
3.2 Datenbereinigung 4
3.3 Orthologie 4
3.3.1 Visual Data Web 5
3.3.2 Datenabhängigkeit 6
3.4 Konvertierung der Daten ins RDF Format 7
3.4.1 Protégé 7
3.4.2 GraphDB 9

4 SPARQL Anfragen **12**
4.1 Diagramme zeichnen 13

5 Poster **16**

6 Lernerfahrungen und Auswertung **16**

7 Zusammenfassung **17**

1 Einleitung

Im Rahmen des Moduls 'Semantik Web' soll sich mit einer Forschungsfrage beschäftigt werden, diese soll anhand von Datenquellen untersucht werden. Dabei sollen mindestens drei unabhängige Datenquellen benutzt werden, welche logisch miteinander verbunden werden sollen. Falls die Daten dabei noch nicht im RDF Form sind, müssen diese erst um konvertiert und in eine Orthologie eingearbeitet werden. Die Forschungsfrage ist folgende: **Besteht eine Korrelation zwischen der zurzeit regierenden Parteien in Deutschland und deren Investitionshöhe in erneuerbaren Energien und wie hat sich der Anteil an erneuerbaren Energien im Zeitraum von 2000 bis 2020 verändert?** . Hierfür werden Daten der Wahl und die der regierenden Parteien sowie die Investition und Anteile der erneuerbaren Energien in Deutschland benötigt.

2 Datenquellen

Um die Forschungsfrage bearbeiten zu können, müssen Daten gefunden werden, welche zu dem Thema passen und aussagekräftig sind. Betrachtet werden Daten im Zeitraum von 2004 bis 2020. Für diesen Zeitraum ist ab 2004 der Aufbau der europäischen Struktur- und Investitionsfonds dokumentiert, hierbei sind auch Daten über Deutschland vorhanden. Da die Bundestagswahl nur aller vier Jahre stattfindet, werden die Zwischenergebnisse durch die Sonntags-frage der Bundestagswahl repräsentiert.

2.1 Verteilung der Erzeugung

'Energy charts' ist eine Seite welche sich mit Themen rund um erneuerbaren Energie beschäftigt (02.07.2021).

Link	https://www.energy-charts.info/charts/energy/chart.htm?l=de&c=DE&stacking=grouped
Datenformat	CSV
Schnittstelle	HTTPS
Lizenz	Creative Commons Attribution 4.0 International
Open Data	⋆⋆⋆

2.2 Wahlen in Deutschland bis 2017

Bundestag.de ist eine offizielle Webseite der Bundesregierung Deutschland, welche sich mit dem Thema rund um die Wahl beschäftigt (02.07.2021).

Link	https://www.bundestag.de/parlament/wahlen/ergebnisse_seit1949-244692
Datenformat	HTML
Schnittstelle	HTTPS
Lizenz	Creative Commons Attribution 4.0 International
Open Data	⋆⋆

2.3 Investition in erneuerbaren Energie bis 2020

Offielle Unter-Seite des Ministerium für Wirtschaft und Energie, welches zur Bundesregierung Deutschland gehört. (02.07.2021).

Link	https://www.erneuerbare-energien.de/EE/Navigation/DE/Service/Erneuerbare_Energien_in_Zahlen/Entwicklung/entwicklung-der-erneuerbaren-energien-in-deutschland.html
Datenformat	Bild
Schnittstelle	HTTPS
Lizenz	Creative Commons Attribution 4.0 International
Open Data	⋆

2.4 Investitionen in Erzeugungsanlagen

Stratista ist für eine Webseite, welche alle möglichen Stratistiken sammeln und teilweise kostenlos anbieten.

Link	https://de.statista.com/statistik/daten/studie/171896/umfrage/investitionen-in-anlagen-zur-nutzung-von-strom-aus-erneuerbaren-energi
Datenformat	CSV
Schnittstelle	HTTPS
Lizenz	Creative Commons Attribution 4.0 International
Open Data	⋆⋆⋆

3 Arbeitsvorgang

Um folgenden soll der Vorgang von den zugrundeliegenden Daten, dessen Suche und Aufbereitung und letztendliche Einarbeitung in das RDF-Schema vorgestellt werden. Hierbei wird auch auf die Erstellung einer eigenen Orthologie auf Grundlagen der Daten eingegangen sowie welche Tools dazu eingesetzt wurden.

3.1 Datensuche

Zu aller erst müssen passende Daten gefunden werden, welche zur Forschungsfrage passen und welche zur Beantwortung dieser Frage beitragen könnten. Hierbei wurde zuerst nach de Wahlergebnissen der Bundesregierung in Deutschland gesucht. Die Bundesregierung besitzt extra ein Archiv, welche genau diese Daten beinhalten, jedoch nur in Tabellenform. Da die Bundesregierung nur aller vier Jahre gewählt wird und zwischenzeitliche keine Daten über die Wahlergebnisse vorhanden sind, müssen die zeitlichen Räume während dieser vier Jahre mit Umfrage-Werten aufgefüllt werden. Hierfür wurden die Daten der Sonntagsfrage der ARD welche monatlich Umfragen über die Wahl der Bundesregierung durchführt. Diese Daten gilt es jetzt nur zu verknüpfen um so Daten in einem vollständigen Zeitraum zu besitzen. Nachdem diese Daten eine Auskunft über die Wahlergebnisse gibt, müssen nun noch Daten über die Investition der Bundesregierung in erneuerbaren Energien und der Gesamtanteil der erneuerbaren Energien gefunden werden. Die Daten, wie viel Investiert wird, konnte auf einer Webseite der Bundesregierung selbst gefunden werden. Leider lagen die Daten nur grafisch vor, sodass diese noch aus dem Bild abgeschrieben werde muss-

ten. Wie hoch nun der Anteil der Erneuerbaren Energien am gesamt Erzeugnisse ist, konnte jedoch nicht in diesen Daten gefunden werden, weshalb hierfür weiter gesucht werden musste. Diese Daten besitzt die Europäische Seite, welche eine vielzahl an Daten anbietet. Somit sind genug Daten vorhanden, um die Forschungsfrage bearbeiten zu können.

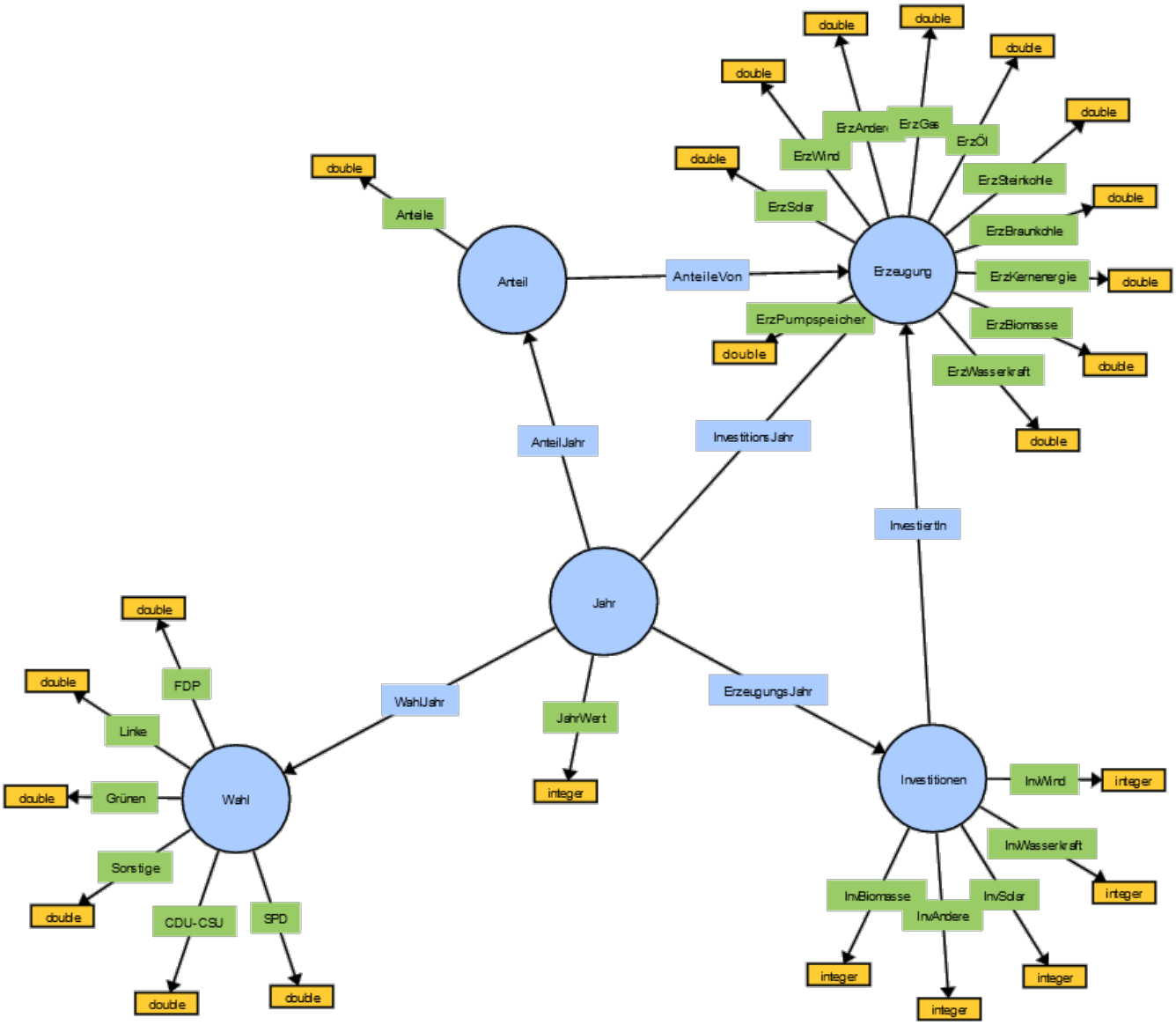
3.2 Datenbereinigung

Um die gefundenen Daten nun einpflegen zu können müssen diese erst bereinigt werden. Da fast alle Daten als Tabelle vorlagen (CSV) musste hierfür nur wenig getan werden, außer die Daten der Investition aus der Grafik in eine Tabelle zu übertragen. Zusätzlich mussten Sonderzeichen oder Umlaute wie ÄÖÜ ersetzt werden, da es sonst zu Problemen im RDF-Schema kommen kann. Da die Daten von der Europäischen Union noch unter anderem Daten über Kohle-kraft und Atomkraft beinhielten mussten diese Spalten aus der Tabelle gelöscht werden, da diese irrelevant für die Forschungsfrage waren. Genauso müssen Leerzeichen, Slashes und Back-Slashes entfernt, oder durch passende Zeichen ersetzt werden. Nachdem die Daten bereinigt wurden, müssen diese noch so umbenannt werden, dass Zusammenhängende Daten auch als solche erkannt werden können. Beispielsweise 'Wasserkaft' hieß in der einen Datenquelle 'Wasserkraft' und in anderen 'Wasserkraft und Wassererzeugnisse', solche müssten auf die selben Wörter unbenannt werden, sodass es möglich ist auch verknüpfte Anfragen zu stellen, welche z.B. Investition und Erzeugungsanteil ausgibt. Auch bei den Daten der Wahlergebnisse und der Wahlprognosen gab es Unterschiede, da sich einzelnen Parteien in Laufe der Zeit unbenannt hatten. Ein weiteres Problem war die Angabe der Zahlen, welche bei der Kommastelle ein Komma besitzt und diese aber im RDF-Format mit einem Punkt gekennzeichnet werden muss, da die Zahl sonst fehlinterpretiert wird, beziehungsweise es zu Fehlern im Importvorgang kommt.

3.3 Orthologie

Da keine der Daten in einer RDF Form sind, muss zunächst eine eigene Orthologie entwickelt und diese Daten dann mit der Orthologie verbunden werden. Zum erstellen der Orthologie wird hierfür wurde das Tool Visual Data Web verwendet. Welche als Webseite zur Verfügung steht und kostenlos nutzbar ist. In Tabelle 1 kann das Vokabular der Orthologie betrachtet werden.

Abbildung 1: Entwickelte Orthologie



3.3.1 Visual Data Web

Visual Data Web biete die Möglichkeit eine Orthologie zu visualisieren, oder eine graphisch zu entwickeln. Um eine Orthologie zu erstellen muss der Bearbeitungs-Modus

Tabelle 1: SPO - Aufbau der Orthologie

Subjekt	**Prädikat**	**Objekt**
Jahr	AnteilJahr	Anteil
Jahr	InvestionJahr	Investitionen
Jahr	ErzeugungsJahr	Erzeugung
Jahr	Wahljahr	Wahl
Investitionen	InvestiertIn	Erzeugung
Anteil	AnteilVon	Erzeugung

aktiv sein. Im Bearbeitungs-Modus können neue Klassen hinzugefügt werden. Dabei ist es auch möglich einer Klasse eine Eigenschaft zuzuweisen und welche Datentyp diese Eigenschaft hat. Wurde die Orthologie fertiggestellt, so ist es möglich diese als Bild oder als JSON-Datei zu exportieren.

3.3.2 Datenabhängigkeit

Alle Daten sollen über das Jahr gefunden und verknüpft werden. Hierfür ist es notwendig, dass die Daten eine Zeiteinheit benötigen, sodass diese mit dem Jahr verknüpft werden können. Da sich die Datenquellen sowie so auf die Zeitliche Veränderungen von Daten beziehen ist es nicht mehr nötig, diese manuell mit einem Zeitpunkt zu versehen. Auf Grundlage der Daten konnte die Die Orthologie erstellt werden, welche mit der Abbildung 1 übereinstimmt. Diese besitzt das Jahr als Schnittstelle für alle anderen Daten, welche mit in der Orthologie enthalten sind. Um die Daten der Wahlergebnisse darstelle zu können wurde eine Klasse 'Wahl' erstellt, welche Jahr und die Anteile der Parteien enthielten und den Typ Zahl hatten. Die Anteile der Erneuerbaren Energien anhand der Gesamt-Erzeugung wurden in einer eignen Klasse gespeichert und nicht selbst berechnet, da sich der Anteil so einfacher abfragen lässt. Die Beträge der Investitionen in erneuerbaren Energien wurden als Double-Wert in der Klasse 'Investition' gespeichert und zusammen mit der Klasse 'Erzeugung' sowie 'Jahr' verknüpft. Die Daten der Erzeugung wurden in der selben namigen Klasse gespeichert und mit 'Anteil', 'Jahr' und 'Investition' verknüpft. Somit konnte eine Orthologie erstellen werden, welche genau zu den Daten passen und eine Abfrage so einfach wie möglich machen sollte.

3.4 Konvertierung der Daten ins RDF Format

Da nun alle Daten bereinigt vorliegen und eine Orthologie als Struktur exisitiert, kann damit begonnen werden die Daten vom CSV Format in das RDF Format zu konvertieren. Dies konnte mithilfe der Programme 'GraphDB' und 'Protégé' umgesetzt werden. Die Anwendung 'Protégé' ist ein kosnteloses quelloffenes Programm, welche für das Erstellen der Semantik benutzt wird. 'GraphDB' ist eine kostenpflichtige Anwendung, welche aber eine Test-Version besitzt und ist für Studenten bis zu 30 Tage kostenlos nutzbar. Dieses Tool soll für die SPARQL-Abfage benutzt werden und benötigt hierfür die Daten in CSV-Format und die Semantik von 'Protégé'.

3.4.1 Protégé

'Protégé' ist eine freie, quelloffene Plattform, die einer wachsenden Nutzergemeinschaft eine Reihe von Werkzeugen zur Verfügung stellt, um Domänenmodelle und wissensbasierte Anwendungen mit Ontologien zu erstellen. Hierfür kann die erstellte Orthologie von 'Visual Data Web' importiert werden. Trotzdem müssen die Klassen nochmals angepasst und die Eigenschaften hinzugefügt werden. Zu Letzt kann die erstellte Struktur expotiert werden, sodass diese in 'GraphDB' benutzt werden kann um die Daten importieren und mit einer SPARQL-Abfrage ausgeben zu können. Im folgenden werden die Arbeitsschritte an einem Beispiel vorgestellt.

Zu aller erst muss für jede Tabelle eine Klasse erstellt werden. Aus den Datenquellen ergeben sich somit fünf Klassen. Diese haben alle die Elternklasse 'Thing', welche die standard Klasse ist und bereits über ein paar Eigenschaften verfügt.

Die angelegten Klassen müssen noch über Eigenschaften miteinander verbunden werden. Diese Eigenschaft soft dafür, dass in den Anfragen später ein korrektes verlinken der Quellen funktioniert.

Zuletzt muss die Orthologie noch exportiert werden. Auch hier gibt es die Möglichkeit verschiedene Datenformate auszuwählen (JSON, TXT, OWL und viele mehr), da aber Graphdb verwendet wird, muss es zwingend das OWL-Format sein. Die W3C Web Ontology Language (OWL) ist eine Sprache des Semantic Web, mit der sich umfangreiches und komplexes Wissen über Dinge, Gruppen von Dingen und Beziehungen zwischen Dingen darstellen lässt. OWL ist eine auf Computerlogik basierende Sprache, so dass das in OWL ausgedrückte Wissen von Computerprogrammen genutzt werden kann, z. B. um die Konsistenz dieses Wissens zu überprüfen oder um implizites Wissen explizit zu machen.

Abbildung 2: Protégé - Klassenansicht

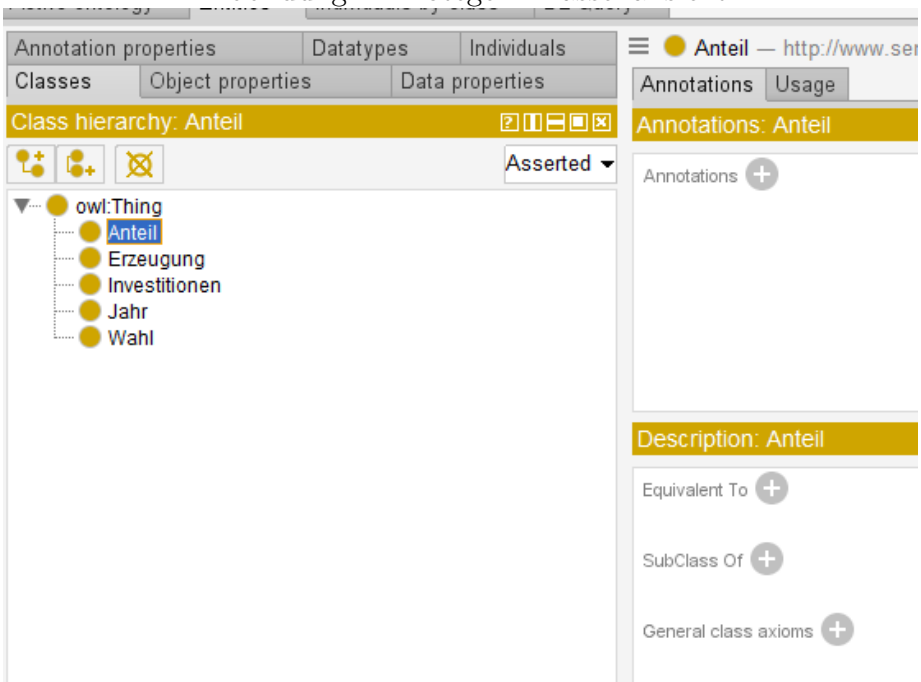
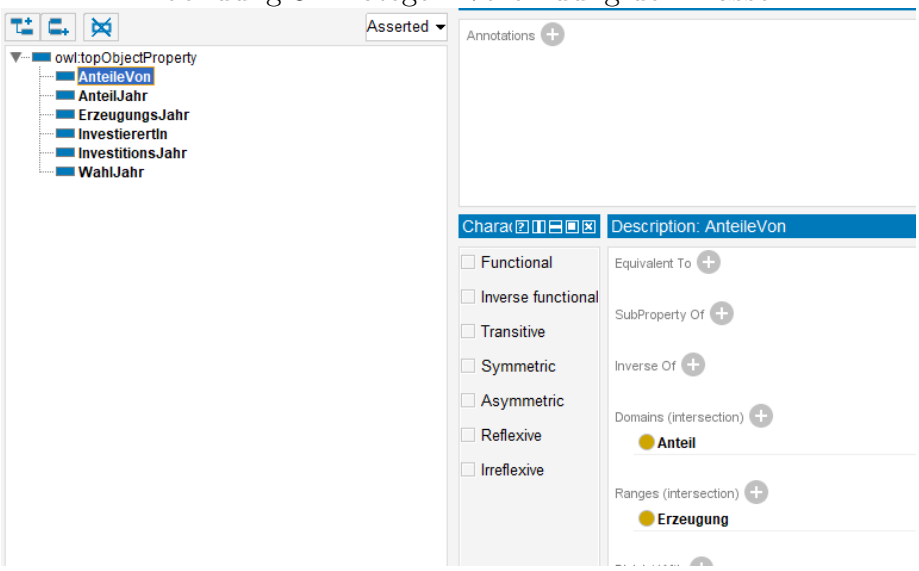


Abbildung 3: Protégé - Verbindung der Klassen

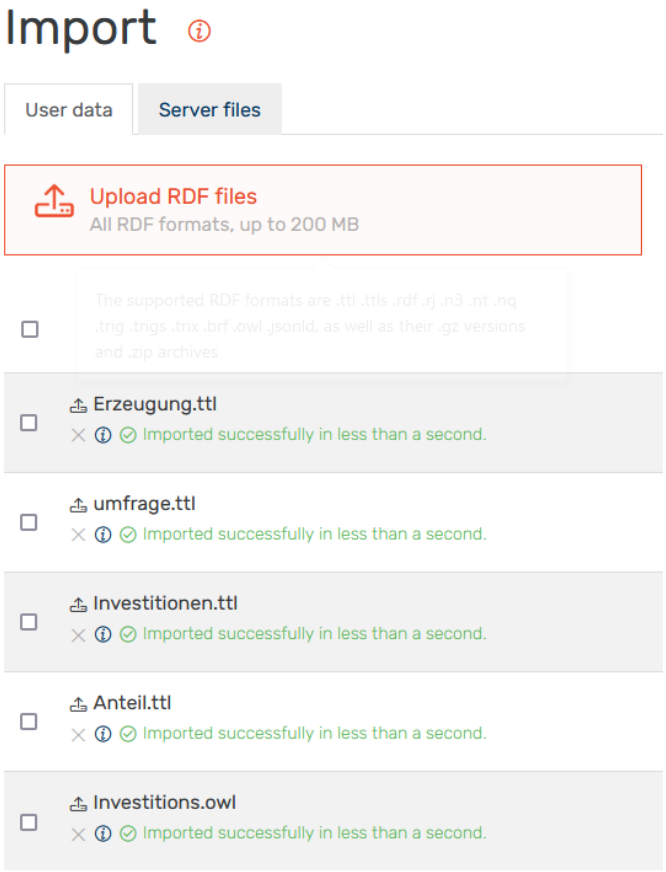


3.4.2 GraphDB

Dieses Kapitel beschreibt die Arbeitsvorgänge mit GraphDB von dem Import der OWL-Datei sowie der Import der Daten aus den CSV-Dateien. 'GraphDB' ist eine Graphen-Datenbank und ein Knowledge-Discovery-Tool, das RDF- und SPARQL-kompatibel ist und als Hochverfügbarkeits-Cluster zur Verfügung steht. 'Ontotext GraphDB' wird in verschiedenen europäischen Forschungsprojekten eingesetzt. Im April 2021 ist Graph DB der viert beliebteste RDF-Speicher und das sechst beliebteste DBMS-System in Europa.

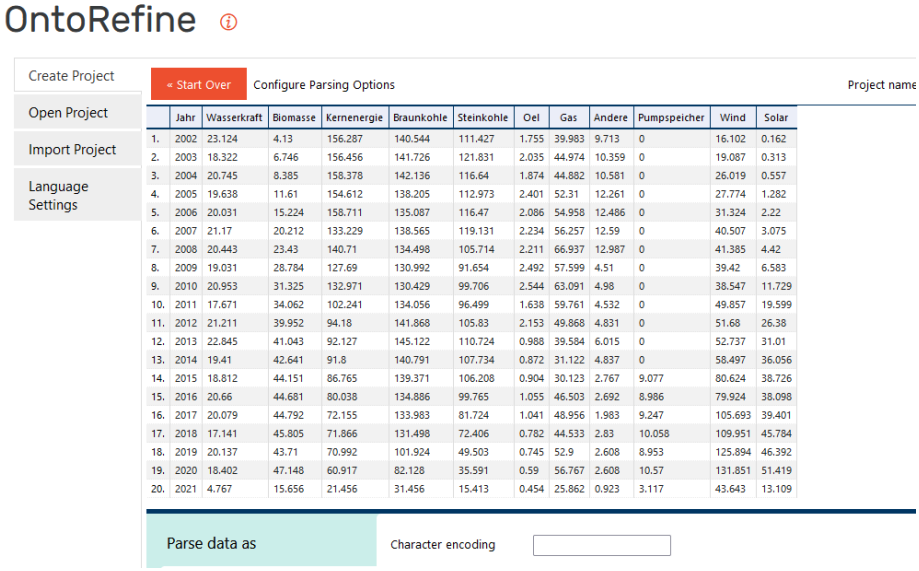
Zunächst muss die OWL-Struktur importiert werde. Dach können die CSV-Daten Importiert werden (siehe Abbildung 4).

Abbildung 4: GraphDB - Import der TTL-Daten



Es wird 'OntoRefine' von 'GraphDB' benutzt um die Daten aus den CSV-Dateien in die Ortholigie zu laden und mit der Sturktur zu verbinden (siehe Abbilding 5).

Abbildung 5: GraphDB - Verbindung der Klassen



	Jahr	Wasserkraft	Biomasse	Kernenergie	Braunkohle	Steinkohle	Oel	Gas	Andere	Pumpspeicher	Wind	Solar
1.	2002	23.124	4.13	156.287	140.544	111.427	1.755	39.983	9.713	0	16.102	0.162
2.	2003	18.322	6.746	156.456	141.726	121.831	2.035	44.974	10.359	0	19.087	0.313
3.	2004	20.745	8.385	158.378	142.136	116.64	1.874	44.882	10.581	0	26.019	0.557
4.	2005	19.638	11.61	154.612	138.205	112.973	2.401	52.31	12.261	0	27.774	1.282
5.	2006	20.031	15.224	158.711	135.087	116.47	2.086	54.958	12.486	0	31.324	2.22
6.	2007	21.17	20.212	133.229	138.565	119.131	2.234	56.257	12.59	0	40.507	3.075
7.	2008	20.443	23.43	140.71	134.498	105.714	2.211	66.937	12.987	0	41.385	4.42
8.	2009	19.031	28.784	127.69	130.992	91.654	2.492	57.599	4.51	0	39.42	6.583
9.	2010	20.953	31.325	132.971	130.429	99.706	2.544	63.091	4.98	0	38.547	11.729
10.	2011	17.671	34.062	102.241	134.056	96.499	1.638	59.761	4.532	0	49.857	19.599
11.	2012	21.211	39.952	94.18	141.868	105.83	2.153	49.868	4.831	0	51.68	26.38
12.	2013	22.845	41.043	92.127	145.122	110.724	0.988	39.584	6.015	0	52.737	31.01
13.	2014	19.41	42.641	91.8	140.791	107.734	0.872	31.122	4.837	0	58.497	36.056
14.	2015	18.812	44.151	86.765	139.371	106.208	0.904	30.123	2.767	9.077	80.624	38.726
15.	2016	20.66	44.681	80.038	134.886	99.765	1.055	46.503	2.692	8.986	79.924	38.098
16.	2017	20.079	44.792	72.155	133.983	81.724	1.041	48.956	1.983	9.247	105.693	39.401
17.	2018	17.141	45.805	71.866	131.498	72.406	0.782	44.533	2.83	10.058	109.951	45.784
18.	2019	20.137	43.71	70.992	101.924	49.503	0.745	52.9	2.608	8.953	125.894	46.392
19.	2020	18.402	47.148	60.917	82.128	35.591	0.59	56.767	2.608	10.57	131.851	51.419
20.	2021	4.767	15.656	21.456	31.456	15.413	0.454	25.862	0.923	3.117	43.643	13.109

Diese müssen nun noch mit der OWL-Struktur verknüpft werden. Die Verlinkung der einzelnen Daten, erfolgt durch den Verweis auf die Spalten, der importieren CSV-Dateien, unter der Verwendung von IRI's. Die Verlinkung der unterschiedlichen Datenquellen erfolgt über die namens gleiche und Inhalt überschneidende Spalte: 'Jahr', welche auf Abbildung 6 zu sehen ist. Dies erfolgt automatisch für Datenquellen mit derselben Basis-IRI mithilfe von GraphDB.

Abbildung 6: GraphDB - Verknüpfung der Elemente

4 SPARQL Anfragen

Folgende Anfragen dienen der Veranschaulichung der Daten, welche nun im RDF-Format vorliegen und ein eigenes Vokabular besitzen. Es sollen zu erst alle Investitionen im Zeitraum von 2004-2020 ausgegeben werden und danach die der Wahl im selben Zeitraum. Zusätzlich sollen die Investition in erneuerbare Energien gegenüber der Erzeugung von erneuerbare Energien, sowie dessen Anteil an der Gesamterzeugung gestellt werden.

```
@prefix owo: <http://www.semanticweb.org/Sebastian/> .

SELECT ?JahrWert ?ErzAndere ?ErzGas ?ErzBraunkohle ?ErzKernenergie
?ErzPumpspeicher ?ErzSolar ?ErzSteinkohle ?ErzWasserkraft
?ErzWind ?ErzOel ?ErzBiomasse ?SPD ?Sonstige as ?SonstigeWaehler
?Linke ?FDP ?Gruenen ?InvAndere ?InvBiomasse ?InvSolar ?InvWasserkraft
?InvWind ?JahrWert

WHERE
{
        ?Jahr owo:AnteilJahr ? Anteil .
        ?Jahr owo:InvestitionsJahr ? Investitionen .
        ?Jahr owo:ErzeugungJahr ? Erzeugung .
        ?Jahr owo:WahlJahr ? Wahl .

        ?Anteil rdf:type owo:Anteil .
        ?Anteil owo:AnteileVon ?Erzeugung .
        ?Investitionen rdf:type owo:Investitionen .
        ?Investitionen owo:InvestitiertIn  ?Erzeugung.
        ?Erzeugung rdf:type owo:Erzeugung .
        ?Wahl rdf:type owo:Wahl .
        ?Jahr rdf:type owo:Jahr .

        FILTER(xsd:integer (?JahrWert) => 2004) .
        FILTER(xsd:integer (?JahrWert) < 2021) .
}

GROUP By ?JahrWert

ORDER By ASC ?JahrWert
```

Als letztes sollen die Daten der Investitionen und die der Wahlen gegenübergestellt werden. Hierbei wird der Zeitraum von 2000 bis 2010 betrachtet. Hierfür müssen nur die anderen Abfrage-Werte entfernt und der Zeitraum angepasst werden.

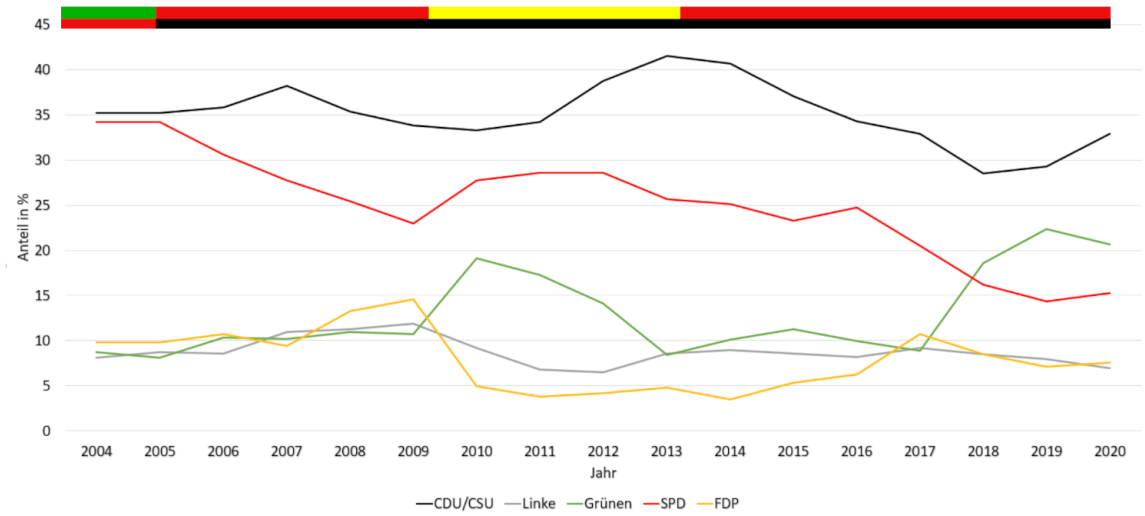
```
@prefix owo: <http://www.semanticweb.org/Sebastian/> .

SELECT ?JahrWert ?SPD ?Linke ?FDP ?Gruenen
?Sonstige as ?SonstigeWaehler ?InvAndere ?InvBiomasse
?InvSolar ?InvWasserkraft ?InvWind ?JahrWert

WHERE
{
        ?Jahr owo:InvestitionsJahr ? Investitionen .
        ?Jahr owo:WahlJahr ? Wahl .

        ?Investitionen rdf:type owo:Investitionen .
        ?Investitionen owo:InvestitiertIn  ?Erzeugung .
        ?Wahl rdf:type owo:Wahl .
        ?Jahr rdf:type owo:Jahr .

        FILTER(xsd:integer (?JahrWert) => 2000) .
        FILTER(xsd:integer (?JahrWert) < 2010) .
}

GROUP By ?JahrWert

ORDER By ASC ?JahrWert
```

4.1 Diagramme zeichnen

Als letztes werden die aus den Anfragen erzeugten Daten noch in Diagramme umgewandelt werden. Dies wurde mit Microsoft Excel bewerkstelligt werden. Hierfür werden die Daten der SPARQL-Anfrage als CSV exportiert und in Excel geladen. Im folgenden werden diese Diagramme kurz vorgestellt. Abbildung 7 zeigt den Wahlverlauf von 2004 bis 2020 und zusätzlich die Gewählte Bundesregierung zu dieser Zeit. Dabei entspricht Grün: Die GRÜNEN, Geld: FDP, Rot: SPD und Schwarz:CDU. Diese Daten wurden nachträglich hinzugefügt, zum besserem Verständnisse.

Das Diagramm von Abbildung 8 visualisiert die Investitionen der Bundesregierung Deutschland von 2004 bis 2019. Es ist deutlich zu Erkennen, dass die Investitio-

Abbildung 7: Wahlverlauf



nen bis 2009 stetig anstiegen aber ab 2009 wieder kontinuierlich fielen. Vor allem die Investition in Solaranlagen sind wieder stark gesunken. Die Beiträge zur Wasserkraft und Biomassen sind relativ zur Windkraft und Solarenergie sehr klein.

Das letzte Diagramm von Abbilding 9 zeigt die Verteilung der Stromerzeugung in Deutschland im Zeitraum von 2004 bis 2020. Es ist zu erkennen, dass die Stein- und Braunkohle sowie Kernenergie über die Jahre hinweg sinken und dafür die Windenergie gestiegen ist. Solarenergie, Wasserkraft und Biomasse sind über den Zeitraum leicht gestiegen.

Abbildung 8: Investitionen in erneuerbare Energien

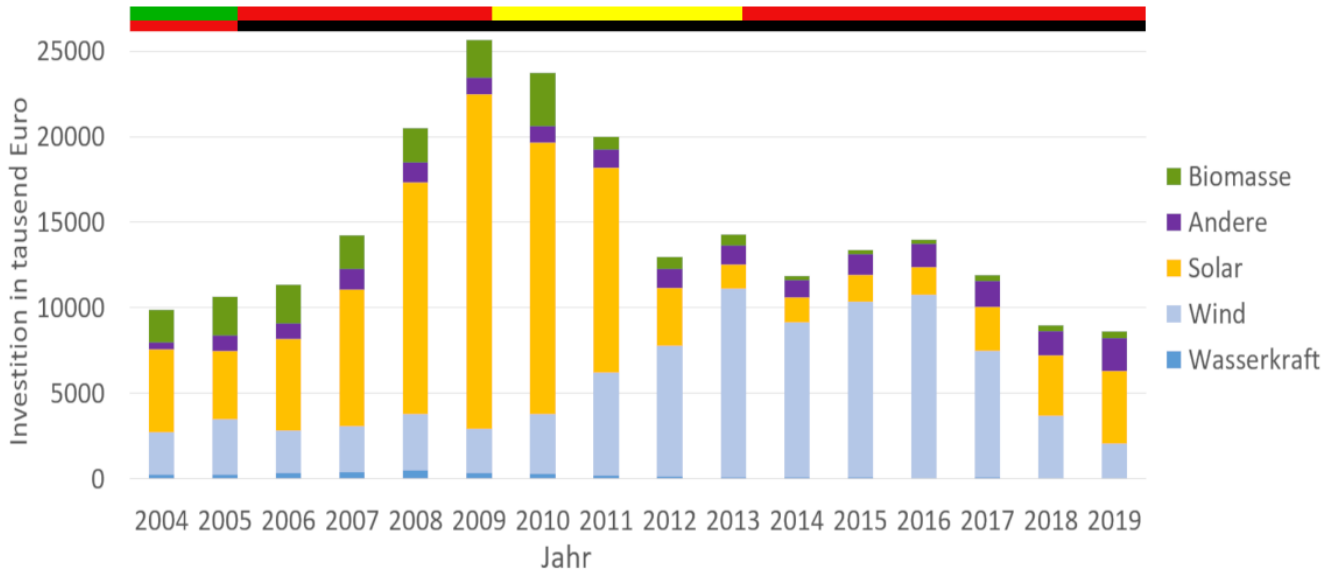
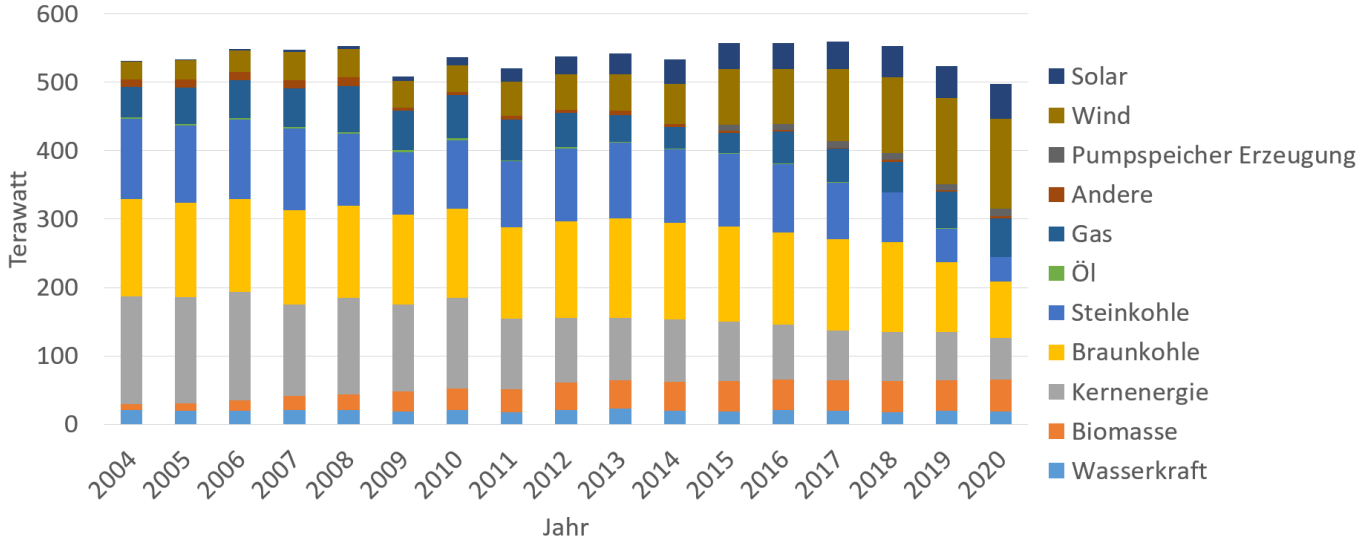


Abbildung 9: Erzeugungsverteilung



5 Poster

Ein Poster sollte zur Vorstellung der Forschungsfrage und des Zwischenstandes erstellt werden. Diese sollte den Titel, die Forschungsfrage, die Orthologie sowie (falls gegeben) die SPARQL-Anfragen. Zusätzlich sollten das Design der HTWK übernommen werden sowie Daten zum Kurs und einen Link für ein mögliches GIT-Repository mit auf das Posten platziert werden. Um die zusammengesammelten Daten bereits vorzustellen wurden die Daten in eine Exceltabelle gespeichert und als Diagramm ausgegeben. Dabei wurden die Wahlergebnisse und die Gesamt-Investitionshöhe in erneuerbare Energien im Zeitlichen verlauf wurden als X-Y Diagramm dargestellt. Zusätzlich wurde über diesen Diagrammen die gewählte Bundesregierung als Farben dargestellt, welche zu der Zeit (Jahr), die Mehrheit hatten.

Es ist schon zu erkennen, dass ab 2009 die Bundesregierung (CDU-SPD) wesentlich weniger Geld in erneuerbare Energien investiert hatte als noch zuvor. In diesem Zeitraum hatten die GRÜNEN auch einen wesentlich höheren Anteil in den Wahlergebnissen.

Somit hat die Bundesregierung es immer mehr der Privatwirtschaft zur Aufgabe gemacht, in erneuerbare Energien zu investieren, da der Anteil der Produktionsmenge zwar stetig steigt, aber die Investitionen stetig sanken. Auch sanken die Investition in Solarkraft und Wasserkraft ab 2009 stark ab. Das Posten kann in Abbildung 10 betrachtet werden.

6 Lernerfahrungen und Auswertung

Bis zu diesem Modul wusste ich gar nicht, dass es RDF bzw. SPARQL Abfragen überhaupt gibt. Da ich viel mit Datenbanken mache und mich gut mit SQL-Abfragen auskenne. Den Vorgang von den Rohdaten zu der SPAQL-Abfrage zu kommen ist leider etwas kompliziert gewesen, da wir in der Vorlesung, dieses Thema erst sehr spät behandelt haben. Trotzdem konnte ich viele neue Dinge in Erfahrung bringen und ich sehe jetzt auch Vorteile gegenüber von SQL-Abfragen, welche immer genaue Tabellenquelle benötigen, nicht wie bei SPAQL-Abfragen, wo dies mit der Orthologie bewerkstelligt wird. Durch dieses Fach konnte ich mir ein grobes Verständnis darüber verschaffen, wie eine RDF-Umwandlung funktioniert und wie man diese Daten dann abfragen kann. Zusätzlich war es nochmals eine gute Übung ein Poster zur Darstellung der Forschungsfrage und der momentanen Werte anzufertigen. Genauso ist die Vorstellung von 'Semantik Web' als Teilgebiet der Informatik, welche sich mit der Verknüpfung von Information beschäftigt, wie auch der Möglichkeit der Lesbarkeit von Informationen durch Software. Auch konnte meine Forschungsfrage gut beant-

worten und habe bemerkt, dass die Bundesregierung viel zu wenig zum Klimaschutz beträgt, als sie eigentlich könnte. Dies kann man auch anhand der Investitionshöhe betrachten, welche seit 2009 sinkt. Es besteht scheinbar ein Zusammenhang von der gewählten Bundesregierung und der Investitionshöhe in erneuerbare Energien. Es sollten dennoch noch weitere Daten in Betracht gezogen werden um diese Annahme zu bestärken.

7 Zusammenfassung

Die Forschungsfrage konnte mit Daten, welche selbständig aus mindestens drei Quellen gesucht und bereinigt wurden, bearbeitet werden. Dabei wurden Tools benutzt, welche die Daten aus den Bild-Daten beziehungsweise CSV-Format in ein RDF-Format konvertierte, womit dann SPAQL-Anfragen gestellt werden konnten. Diese Anfragen wurden dazu benutzt um Diagramme zu zeichnen und möglichst so zu Gestalten, dass diese die Forschungsfrage beantworten. Das Repository kann unter dem Link https://github.com/poetter-sebastian/Semantik-Web erreicht werden. Dieses Projekt beinhaltet alle Stadien von Dateien, also roh-, bearbeitete und nutzbare Dateien. Darüberhinaus diese Dokumentation sowie die des Posters.

Dieses Projekt konnte erfolgreich umgesetzt werden und das Teilgebiet, des 'Semantik Web' konnte gut und lehrreich vorgestellt werden.

Abbildung 10: Fertiges Poster

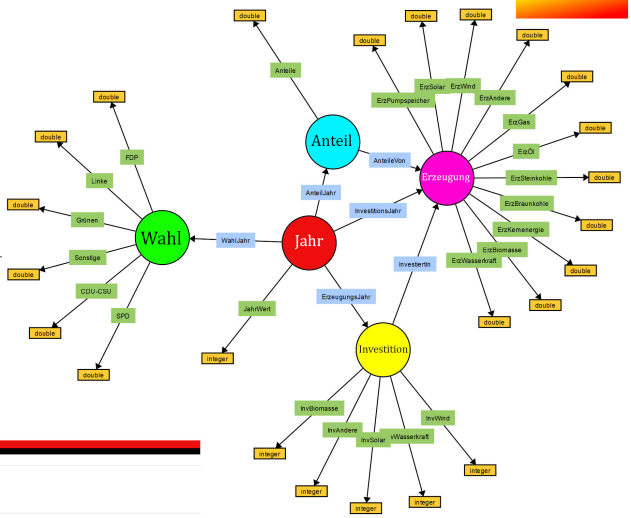
Investitionen in grüne Energie der Bundesregierung Deutschland

Forschungsfrage:

Besteht eine Korrelation zwischen der gewählten Bundesregierung in Deutschland und deren Investitionshöhe in erneuerbaren Energien und wie hat sich der Anteil an erneuerbaren Energien im Zeitraum von 2000 bis 2020 verändert?

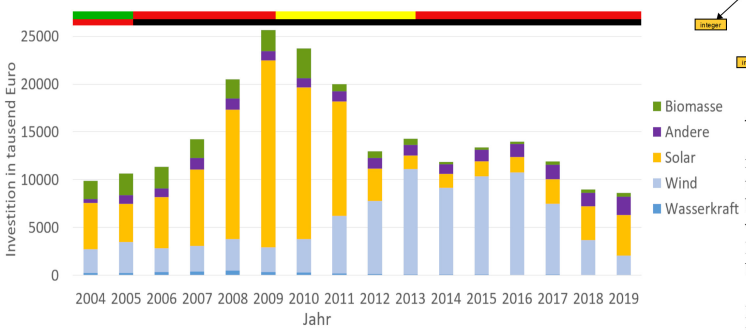
Datenquellen

- **Verteilung der Erzeugung:** https://www.energy-charts.info/charts/-energy/chart.htm?l=de&c=DE&stacking=grouped (02.07.2021)
- **Wahlen in Deutschland bis 2017:** https://www.bundestag.de/-parlament/wahlen/ergebnisse_seit1949-244692 (02.07.2021)
- **Investition in erneuerbaren Energie bis 2020:** https://-www.erneuerbare-energien.de/EE/Navigation/DE/Service/Erneuerbare_Energien_in_Zahlen/-Entwicklung/entwicklung-der-erneuerbaren-energien-in-deutschland.html (02.07.2021)
- **Investitionen in Erzeugungsanlagen:** https://de.statista.com/statistik/daten/studie/171896/umfrage/investitionen-in-anlagen-zur-nutzung-von-strom-aus-erneuerbaren-energien/ (02.07.2021)



Datenzeitraum

Betrachtet werden Daten von 2004 bis 2020. Für diesen Zeitraum ist ab 2020 der ESI dokumentiert. Da die Bundestagswahl nur aller vier Jahre stattfindet, werden die Zwischenergebnisse durch die Sonntags-frage der Bundestagswahl repräsentiert.



Ausblick

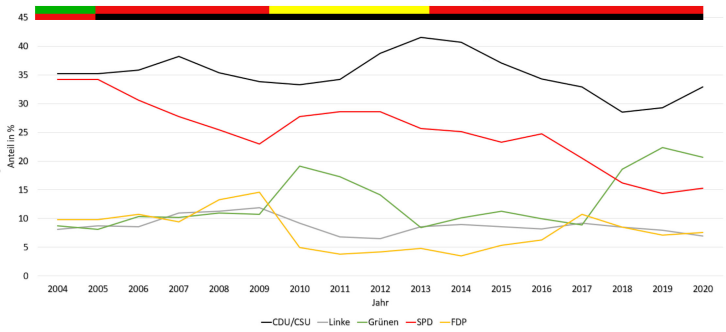
Die Ergebnisse der SPARQL-Anfrage sollen in ihrem Zusammenhang grafisch aufbereitet und analysiert werden um die Frage zu beantworten. Es soll außerdem eingeschätzt werden, in welchem Maß die Wahl der Daten(-quellen) die Beantwortung der Frage beeinflusste.

Die gezeigten Diagramme wurden durch die Daten einer SQL Abfrage erzeugt, da die SPARQL-Anfragen derzeit noch in Arbeit sind.

Auszüge aus den Daten

Es ist schon jetzt zu erkennen, dass ab 2009 die Bundesregierung (CDU-SPD) wesentlich weniger Geld in erneuerbare Energien investiert hat als noch zuvor. In diesem Zeitraum hatten die GRÜNEN auch einen wesentlich höheren Anteil.

Somit hat die Bundesregierung es immer mehr der Privatwirtschaft zur Aufgabe gemacht, in erneuerbare Energien zu investieren, da der Anteil der Produktionsmenge stetig steigt.



I-ITWK
Sebastian Pötter
Semantic Web - 07.07.21
https://github.com/poetter-sebastian/Semantik-Web
IM Fakultät Informatik und Medien