



D6.5

Erstellung eines Gutachtens als Grundlage für einen Ausblick zum benötigten Finanzvolumen für den Produktionsmodus

Inhaltsverzeichnis

1	Einleitung	5
	Zentrale Erkenntnisse	6
1.1	Wissensgraph als Wissensrepräsentation einer Verwaltungsrealität von Nachweisanforderung und Nachweisbereitstellung	6
1.2	Wissensgraph als Lösung bei semantischer Uneindeutigkeit	7
1.3	Wissensgraphen entstehen parallel in verschiedenen Kontexten	11
1.4	Reife der Softwarelandschaft	12
2	Ausbauperspektiven	13
2.1	RegCheck als Datenbackbone eines Verwaltungs-Chatbots	13
2.1.1	Motivation	13
2.1.2	Funktionen des Verwaltungschatbots	14
2.1.3	Zusammenhang zwischen RegCheck 2024 und dem Verwaltungschatbot	17
2.2	RegCheck als Auskunftssystem für die Verwaltung	18
2.2.1	Motivation	18
2.2.2	Verwendung des Wissensgraph durch Mitarbeitende der Verwaltung	19
2.3	RegCheck als Adapter für EU-OOTS Daten und Datenräume der EU	19
3	Pflegekonzept	26
3.1	Pflege des Wissensgraphen und des Datenmodells	26
3.2	Einarbeitung Change Requests	28
3.3	Management der Pflege	29
3.3.1	Regelbetrieb, Support und Beratung	29
3.3.2	Stakeholdermanagement	30
3.3.3	Änderungsmanagement	30
4	Designentscheidungen in der Modellierung	31
4.1	Ontologie und Modellstruktur	32
4.2	DSMeld	37
4.3	FIM	38

4.4	Offene Fragestellungen	40
4.5	Abweichungen von und Änderungsvorschläge zu den EU Core Vocabularies	44
4.5.1	Abweichungen vom Kernmodell CCCEV	44
4.5.2	Änderungsvorschläge an CCCEV	46
5	Kostenkalkulation	50
6	Fazit	52
6.1	Glossar	53
6.2	Abkürzungsverzeichnis	54
7	Anhang	56
7.1	Geklärte Fragestellungen	56

Abbildungsverzeichnis

Abbildung 1: QR-Code zum Projekt	5
Abbildung 2: Use Cases, Spezifikation und Knowledge Graph im Semantic Web Stack	7
Abbildung 3: Mehrsprachiges Konzept als Publications Office Authority Table (target-audience)	10
Abbildung 4: Zitat DSGVO aus Handreichung zur Umsetzung der DSGVO	11
Abbildung 5 Beispiel Unterhaltsvorschuss (Bürgerinnen und Bürger), eigene Darstellung	16
Abbildung 6 Beispiel Gewerbeanmeldung (Unternehmen), eigene Darstellung	17
Abbildung 7: EU OOTS im Semantic Web Stack	21
Abbildung 8: Rolle des Semantic Repositories beim EU OOTS	21
Abbildung 9: EU Datenraum trifft auf nationale Datenraumgrenze	22
Abbildung 10: EU Datenraum durchbricht dank nationalem CCCEV-AP Profil die Datenraumgrenze	22
Abbildung 11: Codelisten vom Publications Office als mehrsprachige URI-versehene RDF-Ressourcen	23
Abbildung 12: beispielhafte Herausforderung von n:n Beziehungen zwischen Providern und Nachweisen im EU OOTS	23
Abbildung 13: Codeliste "Media Type" im FDK 0.4 mit platzsparendem Byte-Kombinationsansatz aus Zeiten des Compilerbaus	24
Abbildung 14: Datenräume der EU, darunter "Öffentliche Verwaltung"	25
Abbildung 15: Prüfkatalogs CAMSS zum Prüfen der Governance und anderer Aspekte vorgeschlagener neuer Vokabulare	29
Abbildung 16: Grundmodell von CCCEV 2.1.0	45
Abbildung 17: Das CCCEV-AP.de (Regcheck) Datenmodell	45
Abbildung 18: RegMo Reifegrad 2.0	46

1 Einleitung

Mit diesem Gutachten soll sowohl Resümee der Projektarbeit gezogen als auch ein Ausblick auf Perspektiven nach dem Projekt gegeben werden.

Dafür werden zunächst die zentralen Erkenntnisse nach knapp zehnmonatiger Projektlaufzeit zusammengefasst. Es wird dargelegt, wie die Technologie eines Wissensgraphen im Kontext der deutschen Verwaltung eingesetzt werden kann. Dabei werden Herausforderungen und Lösungen dargelegt.

Des Weiteren wird darauf eingegangen, in welchem Kontexten Wissensgraphen bereits zum Einsatz kommen und wie sie genutzt werden.

Mit der Beschreibung der Ausbauperspektiven wendet sich der Blick nach vorn. Es wird aufgezeigt und beschrieben, innerhalb welcher Szenarien die erprobte Technologie in der Verwaltungsdigitalisierung zum Einsatz kommen kann. Dabei werden sowohl Bürger und Unternehmen als auch die Sachbearbeiter und der Anschluss zu SDG / EU-OOTS untersucht. In diesem Zusammenhang werden ebenfalls die Aspekte Pflege und Betrieb thematisiert. Ausführlich dargelegt werden die Designentscheidungen hinsichtlich Anpassungen des europäischen Datenmodells an nationale Bedarfe. Final werden Leistungsbausteine für ein Pflichtenheft formuliert und ein Grobausblick auch hinsichtlich benötigter Finanzmittel für eine solches Projekt aufgeschlüsselt. Dabei stellt das Gutachten das Finanzvolumen zur Herstellung des Produktionsmodus unter Berücksichtigung der Inbetriebnahme des Systems sowie der Pflege der Ontologien bzw. Überführung in eine Pflegestelle zusammen.

Kurz nach der Finalisierung wurde eine virtuelle Informationsveranstaltung durchgeführt, deren Aufzeichnung und Folien Sie unter offene-daten.de/Verwaltung bzw. diesen QR-Code nachvollziehen können.



Abbildung 1: QR-Code zum Projekt

Zentrale Erkenntnisse

Im Folgenden werden die zentralen Erkenntnisse dargestellt, die sich im Verlaufe des Projektes ergeben haben. Diese stellen gemeinsam mit dem Abschnitt „Designentscheidungen“ den im Projekt gewonnenen Erkenntnisstand im Hinblick auf die Einsetzbarkeit der Technologie eines Wissensgraphen im Kontext der deutschen Verwaltung dar. Der Abschnitt „Designentscheidungen“ geht dabei detailliert auf getroffene Entscheidungen in der Datenmodellierung ein, während der vorliegende Abschnitt übergeordnete Fragen behandelt.

1.1 Wissensgraph als Wissensrepräsentation einer Verwaltungsrealität von Nachweisanforderung und Nachweisbereitstellung

Zentrales Ziel des Projektes war die Abbildung des im Kontext der Registermodernisierung als vorrangiges Zielbild identifizierten Reifegrads D1 („Datenfeld-Ebene“) im Kontext von Nachweisanforderung und Nachweisbereitstellung. Hierzu wurden zwei Kernvokabulare der EU (Core Criterion and Core Evidence Vocabulary¹; Core Public Service Vocabulary²) vereint und auf den Anwendungszweck zugeschnitten als Profil „CCCEV-AP.de“ veröffentlicht.

Das entstandene Datenmodell konnte die als Ausgangspunkt genommene, über den DS-Meld-Datensatz³ normierte Datenstruktur und die dazugehörigen Verwaltungsverfahren aufnehmen und die Logik von Verwaltungsleistungen, Informationsanforderungen auf Datenfeldebene und Nachweisen auf Datenfeldebene abbilden. Weitere Metadaten, die für die Erreichung des Reifegrads D1 als besonders förderlich erkannt wurden, wurden in Interviews mit Sachbearbeitern testweise erhoben. Durch deren Abbildung in der CCCEV-AP.de-Struktur wurde ein offen im Internet verfügbarer „Wissensgraph der deutschen Verwaltung; hier: Registermodernisierung“ weiterentwickelt.

¹ Github CCCEV <https://github.com/SEMICeu/CCCEV>

² Github CPSV-AP <https://github.com/SEMICeu/CPSV-AP>

³ Versionen des Datensatzes für das Meldewesen bei der KoSIT: <https://www1.osci.de/meldewesen/dsmeld-13316>.

Das Projekt im Semantic Web Stack

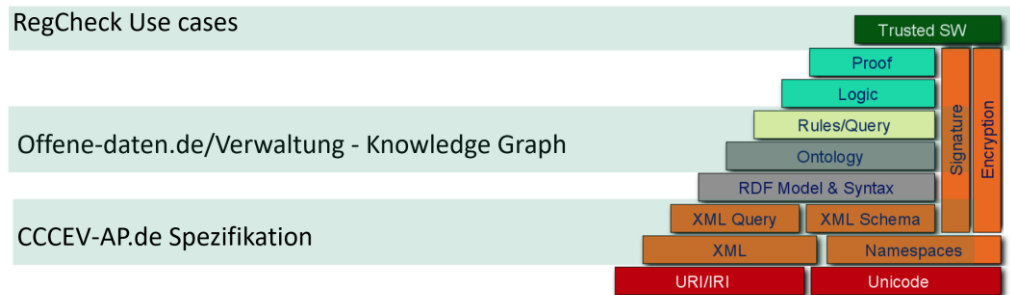


Abbildung 2: Use Cases, Spezifikation und Knowledge Graph im Semantic Web Stack

Die testweise Erweiterung um zwei weitere Fachlichkeiten (Bildungswesen/Immatrikulationsbescheinigung; Finanzwesen/Einkommenssteuerbescheid) konnte ebenfalls im Datenmodell abgebildet werden.

In Summe konnten folgende grundsätzliche Feststellungen getroffen werden:

- Die Datenstruktur ist grundsätzlich in der Lage, beliebige fachliche (horizontale wie vertikale) Erweiterungen abzubilden;
- Der Wissensgraph ist konzeptionell erweiterbar um Attribute oder Beziehungen;
- Die im Wissensgraph hinterlegten Informationen sind zunächst eine Repräsentation von Wissen; sie beschreiben einen Ausschnitt der Realität im Hinblick auf eine Fragestellung („Welche Verwaltungsverfahren benötigen welche Datenfelder?“), definieren damit noch keinen Nutzungszweck für diese Informationen.

Die Frage der Nutzung der im Wissensgraph hinterlegten Informationen wird im Kapitel 2, Ausbauperspektiven behandelt. Das folgende Kapitel widmet sich der Frage, welche Vorteile bei der Nutzung eines Wissensgraphen gegenüber anderen Herangehensweisen im Verlaufe des Projektes offenbar geworden sind.

1.2 Wissensgraph als Lösung bei semantischer Uneindeutigkeit

Die Herausforderung bei der Aufnahme neuer Instanzdaten in das Modell hat sich im Bereich der Semantik gezeigt. Inwiefern Datenfelder mit übergreifenden Konzepten (Namen, Adressdaten) aus verschiedenen Fachlichkeiten, die nicht eindeutig auf den DSMeld-Datensatz rekurrieren, auf dieselbe Entität verweisen können, ist eine fachliche Frage.

Diese Frage konnte durch die bisherigen Normierungen in der deutschen Verwaltungsdigitalisierung nur unzureichend gelöst werden: FIM-Leistungsbeschreibungen⁴ sind textliche und damit vielfach uneindeutige Beschreibungen von erforderlichen Nachweisen. Das Einmelden von FIM-Datenfeldern⁵ aus unterschiedlichen fachlichen Strukturen, dass eine Normierung auf Datenfeld-Ebene erlauben würde, erfordert hohen Abstimmungsbedarf und ist bisher nur in begrenztem Umfang erfolgt.

Das Fachdatenkonzept (FDK) in seiner sich in Abstimmung befindenden Version 0.4 räumt dieser Thematik im Kontext eines geplanten RegMo-Repository ebenfalls großen Raum ein. Die vorliegenden Überlegungen des Fachdatenkonzepts suggerieren, dass Fachdatenkonzept könne die zwischen datenanfordernden Verwaltungsleistungen (nachweisanfordernder Stelle) und datenbereitstellenden Registern (nachweisliefernde Stelle) bestehende Uneindeutigkeit hinsichtlich der Bezeichnung von Datenfeldern auflösen. Die bestehende Systematisierung von Datenfeldern in FIM soll hier primär einfließen.

Vor dem Hintergrund der im Projekt RegCheck gewonnenen Erkenntnisse lassen sich hier folgende Punkte als klärungsbedürftig identifizieren:

- Es kann voraussichtlich keine Homogenität bei den Nachweisbedarfen auf Datenfeld-Ebene zwischen Verwaltungseinheiten auf gleicher Verwaltungsebene angenommen werden (z.B. zwischen zwei benachbarten Kommunen des gleichen Bundeslandes). Die Erkenntnisse der im Projekt geführten Interviews deuten vielmehr darauf hin, dass eine gemeinsame Gesetzesgrundlage durchaus zu einem leicht unterschiedlich gehandhabten Vollzug führen kann. Es erscheint sinnvoll, hier Differenzen herausarbeiten und darstellen zu können, auch um erkennbar zu machen, wo Homogenität besteht (z.B. könnten benachbarte Kommunen gemeinsame länderspezifische Anforderungen haben);
- Die bereits in verschiedenen Fachlichkeiten bestehende semantische Harmonisierung von Konzepten (z.B. „Was ist ein Studienfach?“) sollte die Richtung für eine weitergehende Harmonisierung in der Verwaltungsrealität vorgeben und dementsprechend abbildbar sein;
- Es benötigt eine andere Aussageebene als nur n:n-Beziehungen zu modellieren („Datenfeld X mappt auf Datenfeld Y“). Es werden qualitative Aussagen über Ähnlichkeitsbeziehungen zwischen Konzepten benötigt. Das im FDK-Papier gewählte Beispiel des „Namen“ erscheint hier unterkomplex und berücksichtigt nicht, dass es Überschneidungen von Teilmengen geben kann.
- Die Interviews zeigen auf, dass in der Verwaltungspraxis der dem Vollzug durch gesetzliche Vorgaben eingeräumte Spielraum nicht immer in FIM-Leistungsbeschreibungen abbildbar ist.
- Während der Modellierung fiel auf, dass mit FIM, RLK, VIP und anderen Systeme der Registermodernisierung zwar einige Datenbanken und Repositorien bereitstehen, um Erkenntnisse zu digitalisierten und harmonisierten Nachweisen und Nachweisbedarfen aufzunehmen, es aber noch kein System gibt, um die Heterogenität vor dem

⁴ Zur Erklärung auf digitale-verwaltung.de: <https://www.digitale-verwaltung.de/Webs/DV/DE/onlinezugangsgesetz/ozg-bund/wissen-aus-dem-digitalisierungsprogramm-bund/leistungsbeschreibung/leistungsbeschreibung.html>.

⁵ Zur Erklärung auf fitko.de: https://docs.fitko.de/fim/docs/datenfelder/DF_Allgemeines/.

Harmonisieren fachgerecht abzubilden. Dies wäre jedoch wichtig, um feststellen zu können, wie Once Only erreicht werden kann und wo man sich aktuell befindet.

- Der Wissensgraph-Ansatz bietet zu diesen Fragestellungen Lösungen an. Erstens können Aussagen auf Metadaten-Ebene kontextualisiert werden. Ein geographischer Bezug einer hinterlegten Aussage lässt sich ebenso abbilden, wie die historische Entwicklung (z.B., weil eine Kommune ihre abweichende Anforderung der für eine Verwaltungsleistung benötigten Datenfelder im Laufe der Zeit auf eine bundeslandweite Handhabung harmonisiert – beide Angaben können in versionierter Form abgelegt werden).

Weiterhin können Beziehungen sowohl in andere Fachlichkeiten als auch zu anderen übergreifenden Systematiken (Registerlandkarte⁶, Verwaltungsinformationsplattform⁷, FIM) hergestellt werden. Die im Wissensgraph erstellten Datenfelder entstammen dabei prioritär der rechtlich normierten Standardisierung (DSMeld), danach der fachlich normierten Standardisierung (z.B. XÖV-Standards⁸) und erst zuletzt der individuellen fachlichen Eingabe. Bei eindeutigen Bezügen in andere Systematiken können diese mittels Vergabe von Identifiern abgebildet werden.

Die Beziehungen zwischen diesen Konzepten können mit Ähnlichkeitsaussagen hinterlegt werden. Von der Aussage, zwei „Dinge“ seien „related“ bis hin zur Aussage, ein „Ding“ habe eine „engere“ oder „breitere“ Bedeutung, als ein anderes, lässt sich im Wissensgraph Unschärfe abbilden, die bei FIM oder FDK stets in Eindeutigkeit aufgelöst werden muss: Hier gibt es nur ein, aus der Normierung heraus erarbeitetes „Ding“ (z.B. das FIM-Datenfeld), auf das aus anderen Kontexten heraus gemappt werden muss.


Diese Feststellung lenkt den Blick auf die Frage, an welcher Stelle im Prozess die Klärung der Zuordnung erfolgen muss. Es steht zu vermuten, dass hier FIM und FDK die gemeinsame Herausforderung teilen: die Eindeutigkeit in einem aufwändigen Abstimmungs- und Redaktionsprozess klären zu müssen.

Es kann in diesem Kontext von europäischen Erfahrungen profitiert werden, wo Konzepte z.B. bei der Übersetzung für Richtlinien und Verordnungen in ein „Übersetzungskorpus“ gebracht werden und als Eurovoc-Thesaurus oder Publications Authority tables den Mitgliedstaaten mehrsprachig zur Verfügung steht:

⁶ BVA Registerlandkarte, <https://registerlandkarte.de/>

⁷ DESTATIS Verwaltungsdaten-Informationsplattform, https://www.verwaltungsdaten-informationsplattform.de/DE/Home/home_node.html

⁸ Zur Erklärung auf der Website der KoSIT: <https://www.xoev.de/xoev-4987>.

 Concept

Kinder und Jugendliche AKTUELL

Version: 20240925-0
 Concept URI: <http://publications.europa.eu/resource/authority/target-audience/CHL>

[Inhalt durchsuchen](#)

Definition:
 children and teenagers up to 17 years old – the age group typically corresponding to the pre-school, primary school and secondary school

Letzte Änderung: 23/11/2023
Gültig seit: 2017-10-04
Identifikator: CHL

Labels
BEVORZUGT: Kinder und Jugendliche

Identifiers
PUB_TARGET: 210

Sprachliche Äquivalente
 BG [Децa и подрастващи](#)
 ES [Niños y adolescentes](#)
 CS [Děti a dospívající](#)
 DA [Børn og teenagere](#)
 DE Kinder und Jugendliche
 ET [Lapsed ja teismelised](#)
 EL [Παιδιά και έφηβοι](#)
 EN [Children and teenagers](#)
 FR [Enfants et adolescents](#)
 GA [Leanaí agus déagóirí](#)
 HR [Djeca i maloljetnici](#)
 IT [Bambini e adolescenti](#)
 LV [Bērni un pusaudži](#)
 LT [Vaikai ir paaugliai](#)
 HU [Gyermekek és tizenévesek](#)
 MT [Tfal u adolexxenti](#)
 NL [Kinderen en tieners](#)
 PL [Dzieci i nastolatki](#)
 PT [Crianças e adolescentes](#)
 RO [Copii și adolescenți](#)
 SK [Deti a dospievajúci](#)
 SL [Otroci in najstniki](#)
 FI [Lapset ja teini-ikäiset](#)
 SV [Barn och tonåringar](#)

[Go to asset list](#)
[Dataset details](#)
[Browse content](#)

Abbildung 3: Mehrsprachiges Konzept als Publications Office Authority Table (target-audience)

Der Wissensgraph könnte in diesem Bereich durch einen anderen Ansatz eine Lösung bieten: Informationen würden direkt aus der Sachbearbeitung autonom und dezentral eingepflegt im Rahmen eines geführten Mapping-Prozesses (siehe Pflegekonzept). Dabei würde die Entscheidung über die Zuordnung von Datenfeldern nur dann geprüft, wenn im geführten Prozess kein im rechtlichen oder fachlichen Rahmen bereits normiertes Datenfeld vorliegt.

Weiterhin könnten in diesem Rahmen „Nutzungszwecke“ für Datenfelder erhoben werden, wie dies im Rahmen der geführten Interviews geschehen ist. Damit würde es möglich, nicht nur die Datenfeldbedarfe von Verwaltungsverfahren zu erschließen, sondern zu verstehen, welcher Sachverhalt mit welchen Informationen geprüft wird. Mittels einer einheitlichen Taxonomie von Verwendungszwecken könnten so Ähnlichkeiten auf einer weiteren Ebene identifiziert werden (durch geometrische Nähe von Einträgen mit gleichem Nutzungszweck im Graph). Zusätzlich könnten einerseits Heterogenität dargestellt sowie andererseits Harmonisierung angestoßen werden.

Eine Eigenschaft „Nutzungszweck“ für die Unterstützung des Reifegrad D1 zählt auch auf die rechtlichen Erfordernisse in der Konzeption eines Nachweisdatenabrufes ein: eine Zweckbindung von erhobenen Daten wird auch von der DSGVO gefordert.

Abbildung 4: Zitat DSGVO aus Handreichung zur Umsetzung der DSGVO

(1) Datenerhebung bei der betroffenen Person**Artikel 13 DSGVO (auszugsweise):**

- (1) Werden personenbezogene Daten bei der betroffenen Person erhoben, so teilt der Verantwortliche der betroffenen Person zum Zeitpunkt der Erhebung dieser Daten Folgendes mit:
- a) den Namen und die Kontaktdaten des Verantwortlichen sowie gegebenenfalls seines Vertreters;
 - b) gegebenenfalls die Kontaktdaten des Datenschutzbeauftragten;
 - c) die **Zwecke**, für die die personenbezogenen Daten verarbeitet werden sollen, sowie die Rechtsgrundlage für die Verarbeitung; (...)
 - e) gegebenenfalls die Empfänger oder Kategorien von Empfängern der personenbezogenen Daten und (...)
- (2) Zusätzlich zu den Informationen gemäß Absatz 1 stellt der Verantwortliche der betroffenen Person zum Zeitpunkt der Erhebung dieser Daten folgende weitere Informationen zur Verfügung, die notwendig sind, um eine faire und transparente Verarbeitung zu gewährleisten:
- a) die Dauer, für die die personenbezogenen Daten gespeichert werden oder, falls dies nicht möglich ist, die Kriterien für die Festlegung dieser Dauer;
 - b) das Bestehen eines Rechts auf Auskunft seitens des Verantwortlichen über die betreffenden personenbezogenen Daten sowie auf Berichtigung oder Löschung oder auf Einschränkung der Verarbeitung oder eines Widerspruchsrechts gegen die Verarbeitung sowie des Rechts auf Datenübertragbarkeit;

1.3 Wissensgraphen entstehen parallel in verschiedenen Kontexten

Im Verlauf des Projektes hat sich gezeigt, dass Wissensgraphen aktuell in verschiedenen Bereichen genutzt werden. Konferenzen wie die Semantics⁹ bieten einen Ausblick auf Anwendungsfälle in Wirtschaft und Forschung sowie auf die Relevanz des Themas in Verbindung mit KI-Lösungen. Auch die Abstimmung mit anderen Projekten im Grenzbereich zur öffentlichen Verwaltung wie dem Projekt der Deutschen Zentrale für Tourismus e.V. haben ein Verständnis für den Mehrwert der Technik gebracht, gleichzeitig aber auch den Unterschied zum RegCheck-Projekt deutlich gemacht (die Daten liegen im Falle des Tourismus-Projektes bereits auf Landesebene vor). Ein eindrückliches Werbevideo des Verbands für Tourismus zeigt uns Bürgern das Potential eines Wissensgraphen im Tourismusbereich auf¹⁰.

Erkenntnisse über Nutzungen in der Verwaltung sind zwar weniger öffentlich präsent, haben sich aber dennoch ergeben. Hier ist zum einen das Beispiel des openDVA-Projektes¹¹ zu nennen. Das Projekt ist umfassender konzipiert durch die Berücksichtigung der für RegCheck nicht relevanten Normenanalyse und hat eine starke Bindung an FIM zur semantischen

⁹ Proceedings der Konferenz 2024: <https://2024-eu.semantics.cc/page/proceedings>.

Proceedings der SEMANTICS 2024 Conference in Amsterdam, <https://ebooks.iospress.nl/volume/knowledge-graphs-in-the-age-of-language-models-and-neuro-symbolic-ai-proceedings-of-the-20th-international-conference-on-semantic-systems>

¹⁰ Open-data-Germany.org Deutsche Zentrale für Tourismus e.V., <https://youtu.be/jYbnydjB7g4>

¹¹ Website zum 2. openDVA-Kongress: <https://indico.rz.uni-jena.de/event/123/>.

Normierung. Zum Ende der Projektlaufzeit trat das vom Bundesfinanzministerium initiierte Projekt,¹² das zur Klärung in der Rechtsetzung uneinheitlich verwendeter, zentraler Begriffe wie „Einkommen“ und „Kind“ ebenfalls einen Wissensgraph aufsetzt, in den Fokus der Betrachtung. Auch dort ist der Ansatzpunkt eher an der Schnittstelle zwischen Normenanalyse und Nachweisdaten Anforderung zu sehen. Beide Projekte enden 2024.

Insbesondere mit letztgenanntem Projekt könnte sich Potential für Synergien ergeben. Dieses hat mit Blick von „oben“ auf Normenanalyse und Rechtsbegriffe eventuell Verwendung für die in RegCheck von „unten“ auf Datenfeldebene erhobenen Informationen und die dabei gewonnenen Erkenntnisse.

1.4 Reife der Softwarelandschaft

Im Hinblick auf den erforderlichen Software-Stack waren drei zentrale funktionale Anforderungen zu Beginn des Projektes festgelegt worden:

- Wissensgraph entwickeln und bearbeiten können
- Wissensgraphen visualisieren und Abfragen darauf fahren können
- Tripelstores für die effiziente Speicherung und Verwaltung des Wissensgraphen

Hinzu kam die Anforderung, eine Validierbarkeit der Instanzdaten zu gewährleisten.

In Konsequenz wurde der Wissensgraph als Turtle-RDF und eine SHACL-Datei für die Validierung (Struktur; verpflichtende properties) erstellt.

Dieses Vorgehen war dadurch bedingt, dass der Wissensgraph quelloffen und nachnutzbar sein sollte. Wie sich in der Softwareanalyse im Verlaufe des Projektes gezeigt hat, waren viele der verbreiteten Lösungen für das Anforderungsprofil nicht passend, so dass ein konkreter Vorschlag für einen Softwarestack nicht erfolgen konnte. Einige der Lösungen bieten einen umfassenden, die o.g. Anforderungen gesamtheitlich abdeckenden Funktionsumfang. Die hier zunächst in Aussicht genommene Lösung Stardog zeigte jedoch stellvertretend für All-In-One-Lösungen, dass hier vielfach Kompromisse zugunsten eines geschlossenen Ökosystems einzugehen sind.

Weiterhin hat sich gezeigt, dass für die Analyse der Daten durch Dritte, sowie auch für die Dateneingabe in den Wissensgraphen, Frontend-Lösungen nötig sind, welche in Handhabung und Visualisierung ohne Fachkenntnisse, möglichst mit deutscher Oberfläche, im Bereich Semantic Web bereits nutzbar sind. Hier steht in Frage, inwiefern bisher nicht vertieft untersuchte kleinere Tools oder Eigenentwicklungen sinnvoll wären.

¹² Darstellung des Projektes auf der 13. XÖV-Konferenz: <https://www.xoev.de/sixcms/media.php/13/Once-Only%20%253D%2520Einander%2520verstehen%2520%2528k%25C3%25B6nnen%2520und%2520wollen%2529.pdf>.

2 Ausbauperspektiven

Der in RegCheck 2024 erarbeitete Wissensgraph ist in der Lage, Zusammenhänge zwischen Verwaltungsleistungen und benötigten Nachweisen (sowohl auf Dokumenten-, als auch auf Datenfeldebene) darzustellen. Dabei kann er komplexe Abhängigkeiten darstellen und prinzipiell um viele weitere Informationen – beispielsweise zu den Verwaltungsleistungen – erweitert werden. Im Gegensatz zu relationalen Datenbanken ist er in erster Linie eine Wissensrepräsentation, die nicht auf einen Nutzungszweck hin konzipiert wurde. Der im Wissensgraphen dargestellte „Ausschnitt der Realität“ ist somit nutzbar für verschiedene Kontexte, in denen die im Wissensgraph abgebildeten Informationen relevant sind.

Im vorliegenden Kapitel werden drei mögliche Einsatzszenarien vorgestellt:

1. Nutzung durch Bürger und Unternehmen am Beispiel eines Chatbots
2. Nutzung durch die Verwaltung
3. Anschluss SDG / EU OOTS

Die genannten Szenarien bilden dabei keine abschließende Liste, sondern sind vielmehr naheliegende, erste Überlegungen.

2.1 RegCheck als Datenbackbone eines Verwaltungs-Chatbots

Anhand des Beispiels eines Chatbots wird beschrieben, wie die in RegCheck 2024 erarbeiteten Informationen konkret für Bürger und Unternehmen nutzbar gemacht werden können. Ziel eines Folgeprojektes wäre es, diese Informationen für Bürger und Unternehmen verfügbar zu machen, sodass ein möglichst großer gesellschaftlicher Mehrwert generiert wird.

2.1.1 Motivation

Bürger sowie Unternehmen sehen es oftmals als Herausforderung sich in der als zu bürokratisch wahrgenommenen deutschen Behördenlandschaft¹³ zu orientieren. Sie haben jedoch viele Anlässe, bei welchen sie für die Erbringung von Leistungen mit der Verwaltung interagieren müssen. Sogenannte Lebensereignisse erfordern eine Interaktion mit der deutschen Verwaltung und bergen die Herausforderung, die richtigen Dokumente für den vorgeplanten Behördentermin dabei zu haben. Es gilt hierbei effizient und zielgerichtet alle Informationen vorab einzuholen, die für die Verwaltungsleistung notwendig sind. Dabei kann es sich um Informationen zur Leistung an sich wie z.B. Kosten, Bearbeitungsdauer oder benötigte

¹³ <https://www.tagesschau.de/inland/belastung-buerokratie-rekord-100.html>.

Nachweise handeln. Es können aber auch Informationen zur Beantragung wie z.B. Online-Dienste, herunterladbare oder vor Ort auszufüllende Formulare oder Informationen zur zuständigen Stelle wie z.B. Kontaktdaten, Öffnungszeiten, Adresse nötig sein. Während diese Informationen zum großen Teil in den Redaktionssystemen der Bundesländer auffindbar sind, sind diese Seiten oft wenig bei Bürgern und Unternehmen bekannt und liefern auf Basis der FIM-Leistungsbeschreibungen nicht immer die Informationen in der notwendigen Präzision.

Um hier Abhilfe zu schaffen, könnte auf Basis der in RegCheck 2024 erarbeiteten Ergebnisse ein Verwaltungschatbot aufgebaut werden, der genau an dieser Stelle ansetzt: Informationen zu Verwaltungsdienstleistungen schnell und nutzerfreundlich zur Verfügung zu stellen und dabei zielgerichtet auf die Bedürfnisse und Gegebenheiten der User einzugehen. Ein weiterer Vorteil ist, dass bereits viele Personen mit einem Chatbot kommuniziert haben und so die Komplexität von Verwaltungsverfahren in einem natürlich wirkenden Gespräch heruntergebrochen werden kann. Ein weiterer Vorteil ist, dass dadurch die sogenannte „Offline-Lücke“ geschlossen werden kann: bisherige Digitalisierungs-Bemühungen setzen auf Online-Nachweise, deren Verfügbarkeit und Auslieferung via NOOTS oder EU-OOTS. Dabei wird von Nutzerzentrierung gesprochen und in OZG 1.0 die einer Ende-zu-Ende Digitalisierung vorgelagerte Antragsdigitalisierung forciert. Auf der Strecke jedoch bleiben bei komplexeren Verwaltungsleistungen in Wirklichkeit die Nutzer: denn eine Mischung aus Online- und Offline vorhandenen Nachweisen sind beizubringen. Der um „offline verfügbare Nachweise“ erweiterte Wissensgraph füllt hier eine aus Usability-Sicht methodische Lücke der Registermodernisierung. Informationen aus der analogen und der virtuellen Welt werden im Wissensgraph verknüpft.

2.1.2 Funktionen des Verwaltungschatbots

Der Verwaltungschatbot könnte die zentrale Anlaufstelle für Personen mit Fragen zu Verwaltungsleistungen werden und möglichst alle dieser Fragen in einem Chat-Gespräch beantworten. Dazu zählen unter anderem:

1. **Nachweise für die Verwaltungsleistungen:** Der Verwaltungschatbot kann dem User mitteilen, welche Nachweise für die Verwaltungsleistung notwendig sind und das in einem detaillierteren Grad als in den FIM-Leistungsbeschreibungen, auf den die bestehenden Redaktionssysteme aufgebaut sind. So kann der Verwaltungschatbot die individuelle Situation des Users berücksichtigen: Ein deutscher Gründer einer GbR benötigt beispielsweise andere Nachweise für die Anmeldung eines Gewerbes als ein türkischer Gründer einer GmbH. Der Verwaltungschatbot kann im Dialog die individuellen Gegebenheiten des Users ermitteln und darauf aufbauend passgenau die Nachweisbedarfe liefern.
2. **Once-Only-Fähigkeit der Nachweise darstellen:** Der User erhält durch den Verwaltungschatbot die Information, welche Nachweise bereits automatisiert abgerufen werden können und daher vom User nicht mehr separat beantragt werden müssen. Dazu wird eine kurze Erklärung ausgegeben, die den automatisierten Abruf beschreibt.

3. Informationen zum Einholen von Nachweisen: Zu den Nachweisen, die noch nicht automatisiert abgerufen werden können, gibt der Verwaltungschatbot weiterführende Informationen und Verlinkungen dazu, wie die Nachweise (bis hin zu Öffnungszeiten und Terminvereinbarungen) durch den User eingeholt werden können.
4. Art der Beantragung: Der Verwaltungschatbot kann in dem Dialog Informationen zur Art und Weise der Beantragung geben: Gibt es einen Online-Dienst und wie ist dieser zu finden? Gibt es eine PDF-Datei und wo ist diese zu finden? Ist der Antrag nur vor Ort oder postalisch durchzuführen? All diese Fragen und weitere kann der Verwaltungschatbot zur Art der Beantragung beantworten.
5. Informationen aus FIM-Leistungsbeschreibungen: Weitere Informationen aus den FIM-Leistungsbeschreibungen bzw. den landesspezifischen Anpassungen dieser könnte der Chatbot ausgeben. Das sind z.B. Kosten für die Verwaltungsleistungen, die Kontakt- und Adressdaten der zuständigen Stelle, Rechtsgrundlagen und Voraussetzungen.
6. Leichte Sprache und Mehrsprachigkeit: Dem User ist es möglich, Anfragen in weiteren Sprachen außer deutsch sowie in leichter Sprache zu stellen und diese durch den Verwaltungschatbot beantworten zu lassen. Damit kann der Chatbot auch den massiv gestiegenen Beratungsbedarf durch zugezogene Personen wie Flüchtlinge, Asylbewerber und Migranten teilweise sogar in ihrer Muttersprache eingehen.

Neben diesen aufgeführten Funktionalitäten sind noch weitere denkbar, da die Informationsbasis des Verwaltungschatbots kontinuierlich erweitert werden kann. Der Funktionsumfang ist nicht als determiniert oder vollständig zu betrachten, er soll lediglich Veranschaulichen, wie die User unterstützt werden können. Ziel ist es, den Usern umfassende Informationen zu Verwaltung und Verwaltungsleistungen bereitzustellen.

In einem Beispiel wird ein fiktiver, aber realitätsnaher Anwendungsfall anhand einer Person dargestellt, die Unterhaltsvorschuss beantragen möchte.

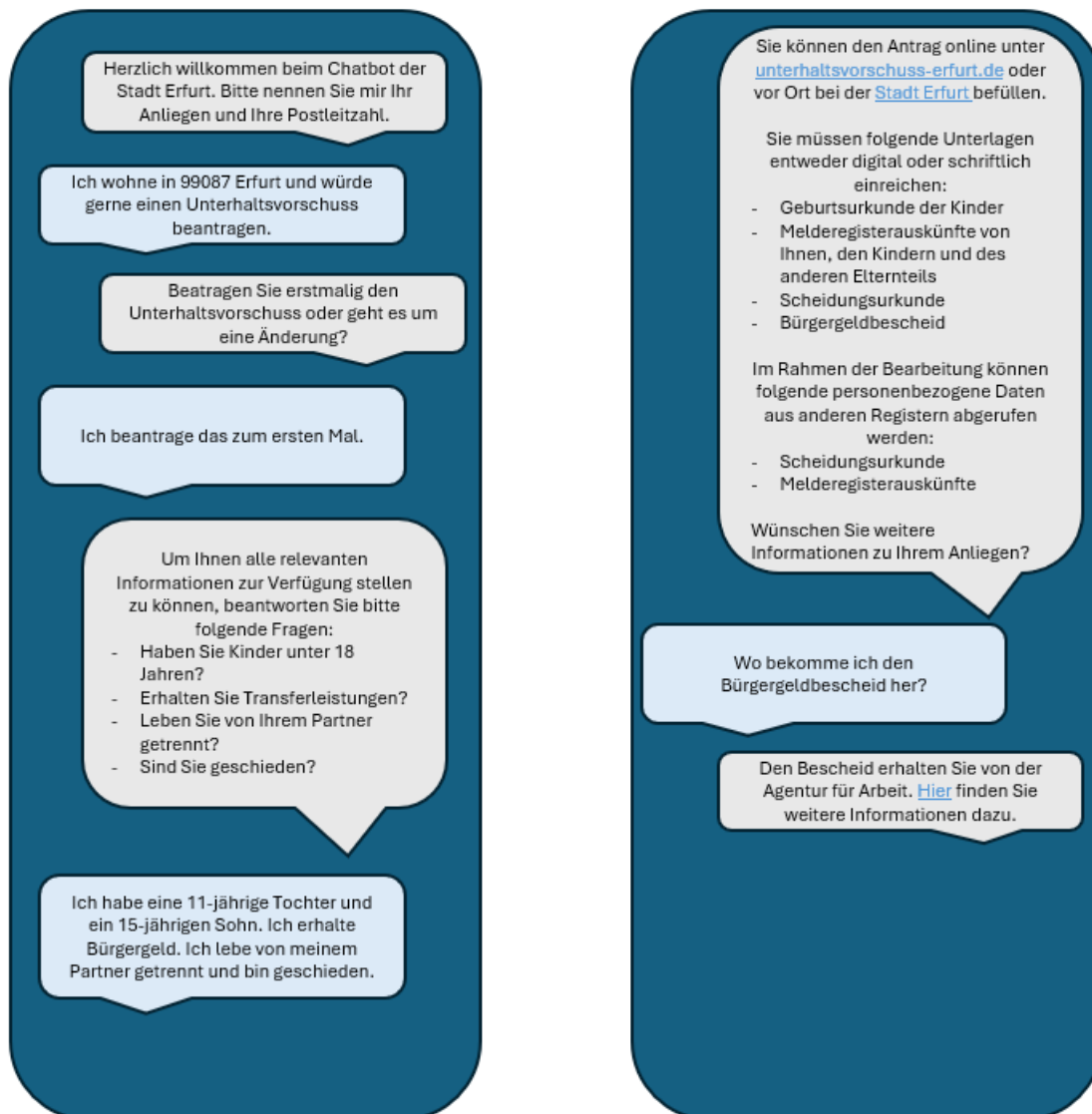


Abbildung 5 Beispiel Unterhaltsvorschuss (Bürgerinnen und Bürger), eigene Darstellung

Der Verwaltungschatbot kann hier mit Antworten unterstützen, die passgenau auf die Situation des Users zugeschnitten sind. Durch Rückfragen wie z.B. der Frage nach Transferleistungen kann das System erkennen, welche Nachweise benötigt werden. Des Weiteren kann der Verwaltungschatbot auf Nachfragen antworten und auch weiterführende Informationen mitteilen.

In einem zweiten Beispiel wird die Interaktion mit einem User dargestellt, der ein Gewerbe anmelden möchte.

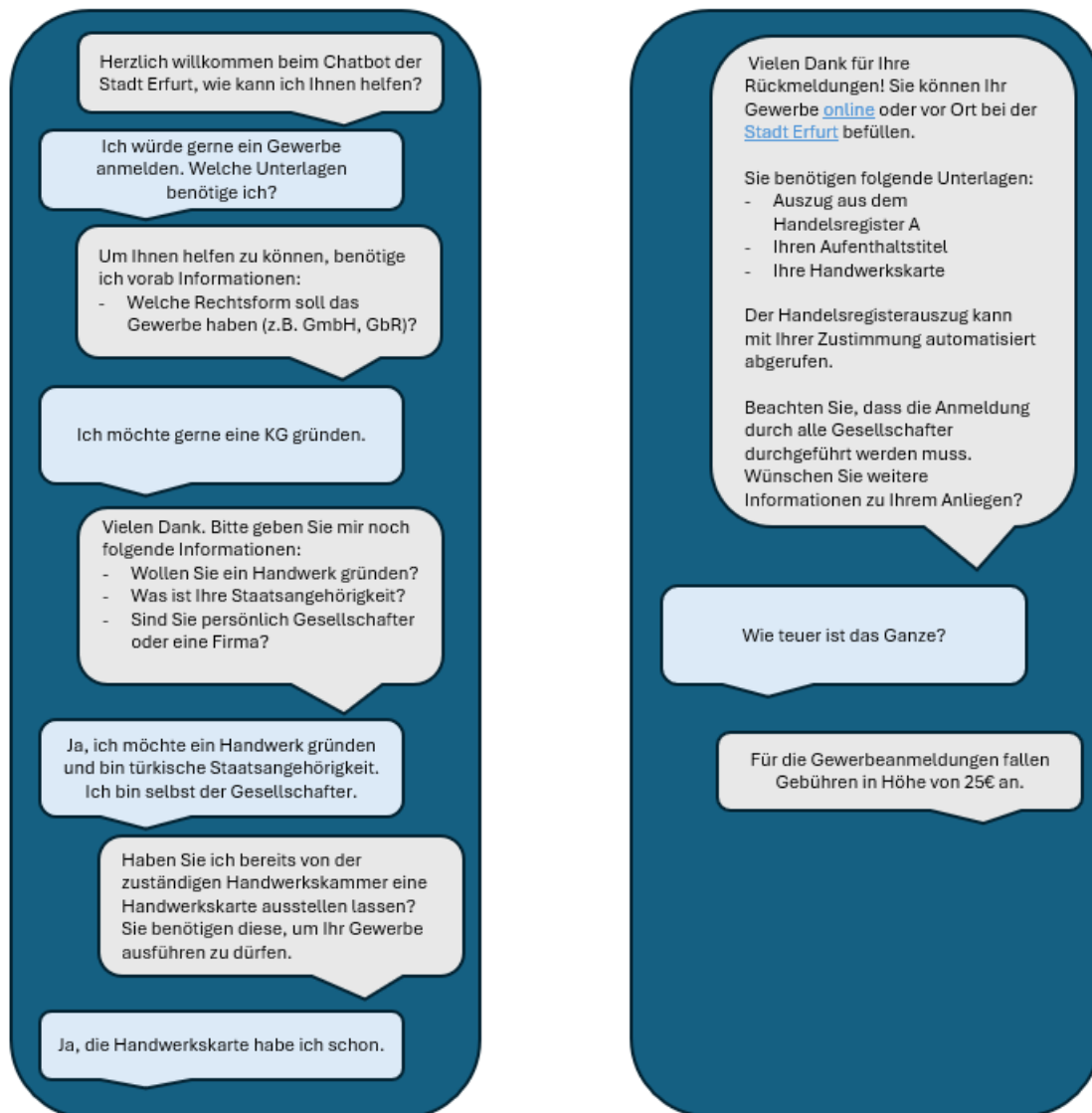


Abbildung 6 Beispiel Gewerbeanmeldung (Unternehmen), eigene Darstellung

Bei einer Gewerbeanmeldung hat unter anderem die gewählte Rechtsform einen großen Einfluss. So divergieren beispielsweise die benötigten Nachweise, je nachdem ob eine Person ein (eingetragenes / nicht eingetragenes) Einzelunternehmen, eine juristische Person oder eine Personengesellschaften darstellt. Der Verwaltungschatbot kann dies erkennen und passgenau die korrekten Nachweise nennen. Er stellt darüber hinaus noch weitere Informationen zur Verfügung.

2.1.3 Zusammenhang zwischen RegCheck 2024 und dem Verwaltungschatbot

Nachdem in 2.1.2 mögliche Funktionen des Verwaltungschatbots skizziert wurden, wird nun erläutert, welcher Zusammenhang zwischen einem möglichen Verwaltungschatbot und den in RegCheck 2024 erarbeiteten Ergebnisse besteht.

Ein auf künstlicher Intelligenz und Natural Language Processing basierender Chatbot benötigt eine gute Datengrundlage, um hilfreiche Antworten für den User geben zu können. Bereits in Kapitel 6.1 des Liefergegenstandes „Künstliche Intelligenz und Registermodernisierung – Potentialanalyse“ werden die Vorteile von Wissensgraphen – wie in RegCheck 2024 entstanden – als Wissensbasis für KI beleuchtet. Da dies in dem Kapitel schon umfangreich behandelt wird, folgt an dieser Stelle nur eine kurze Zusammenfassung:

Ein Wissensgraph ist eine empfehlenswerte Grundlage für künstliche Intelligenz, da die Daten semantisch verknüpft sind und die Bedeutung für Schlussfolgerungen und kontextuelles Verständnis verfügbar gemacht wird. Außerdem ist das System als „offene Welt“ konzipiert, bei der nicht erst alle Ausprägungen vorliegen müssen, um Daten in einem Datenmodell ablegen zu können. Dies ermöglicht, dass KI-Systeme durch den Wissensgraph neue Erkenntnisse ableiten sowie verschiedene Datenquellen nahtlos miteinander kombinieren können. Nutzer können Rückmeldung geben und so ein „Lernende System“ unterstützen, bei dem eher unpassende Nachweis-Zuordnungen sich „nach und nach aus dem System herauswachsen“.

Des Weiteren ist der Wissensgraph als Datengrundlage problemlos erweiterbar und neue Informationen können stets integriert werden. Außerdem wird die Erklärbarkeit der Antworten sowie die Transparenz erhöht, wodurch der Output der KI nachvollziehbar wird.

2.2 RegCheck als Auskunftssystem für die Verwaltung

Neben den Bürgern und Unternehmen kann auch die Verwaltung von dem im Wissensgraphen modellierten Informationen profitieren. Szenarien zur Nutzung durch die Verwaltung wurde bereits in Liefergegenstand D3.2 anhand der Persona-Methode beleuchtet. Da dieser Liefergegenstand den Blick auf zukünftige Verwendungen wirft, werden diese Ansätze hier aufgegriffen.

2.2.1 Motivation

Nicht nur Bürger und Unternehmen sind von der Komplexität der deutschen Verwaltung betroffen, sondern auch die Personen, die in der Verwaltung arbeiten und sich an der Digitalisierung dieser beteiligen. Anträge, Rechtsgrundlagen, Nachweise, Register und Online-Dienste stehen im engen Zusammenhang miteinander und ergeben eine enorme Vielschichtigkeit, insbesondere da die Anzahl der Gesetze, Verordnungen und Einzelnormen stetig steigt.¹⁴ Damit die Mitarbeitenden der Verwaltung den Überblick bewahren können und Zusammenhänge ohne großen Zeitaufwand prüfen können, kann der Wissensgraph von RegCheck verwendet werden.

¹⁴ <https://www.tagesschau.de/inland/gesellschaft/buerokratie-regelungsdichte-zunahme-100.html>.

2.2.2 Verwendung des Wissensgraph durch Mitarbeitende der Verwaltung

Auch die Mitarbeitenden der Verwaltung können von dem Wissensgraphen profitieren. Voraussetzung dafür ist zum einen, dass der Wissensgraph noch maßgeblich erweitert wird und zum anderen, dass der Wissensgraph nutzerfreundlich einsetzbar gemacht wird. Die Pflege wird bereits in Kapitel 3 dieses Liefergegenstandes ausführlich dokumentiert. Die hinterlegten Informationen müssen durch ein Frontend-Interface abrufbar sein, das ohne Kenntnisse von Semantic-Web-Techniken nutzbar ist. Erste Ansätze dazu, wie die Oberflächen in Form eines X-Wikis,¹⁵ wurde diskutiert. Entsprechend muss auch Software entwickelt oder nachgenutzt werden, die die Informationen des Wissensgraphen zielgruppengerecht verfügbar macht. Dafür sind verschiedene Ausprägungen denkbar.

Grundsätzlich kann der Wissensgraph in weiteren Anwendungsgebieten eingesetzt werden, darunter die Qualitätssicherung von FIM-Datenfeldern, bei der Erstellung oder Anpassung von Rechtsgrundlagen zu Nachweisen und Nachweisbedarfen, bei der Weiterentwicklung von Online-Diensten nach Nachweisabrufreiheitsgrad D1, sowie bei der Erstellung und Bearbeitung von Anträgen für Verwaltungsleistungen. In jedem dieser Anwendungsfälle hilft der Wissensgraph dabei, komplexe Datenbeziehungen und Abhängigkeiten verständlich und zugänglich zu machen. Die Fähigkeit des Wissensgraphen, relevante Daten und Zusammenhänge, beispielsweise durch die Identifikation passender Datenfelder, darzustellen, kann für die Analyse von Auswirkungen von Rechtsänderungen auf Nachweise und Verwaltungsverfahren oder bei der Unterstützung des datensparsamen Nachweisabrufes von großem Wert sein. Tiefergehende Informationen zu den Anwendungsfällen sind in Liefergegenstand D3.2 zu finden.

2.3 RegCheck als Adapter für EU-OOTS Daten und Datenräume der EU

Im Kontext der Umsetzung der Single-Digital-Gateway-Verordnung¹⁶ der EU zum grenzüberschreitenden Nachweisabruf zwischen EU-Mitgliedsstaaten im Hinblick auf Nachweise für eine Reihe spezifischer Verfahren (Anhang II der Verordnung EU 2018/1724, siehe Fußnote 16) finden aktuell auf nationaler Ebene Abstimmungsprozesse hinsichtlich der hiervon betroffenen „Verwaltungsverfahren, Online-Services, Nachweise, deren Inhalte und Register“ statt.¹⁷

Für den geplanten grenzüberschreitenden Nachweisabruf hat die EU-KOM folgende Infrastruktur zur Verfügung gestellt:

¹⁵ XWiki- Frontend Möglichkeit ergänzend zum Wissensgraph, <https://www.xwiki.org/xwiki/bin/view/Main/WebHome>

¹⁶ <https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32018R1724>

¹⁷ https://www.digitale-verwaltung.de/SharedDocs/downloads/Webs/DV/DE/sdg-factsheet-evidence-survey.pdf?__blob=publicationFile&v=7

1. Semantic Repository¹⁸
Ein zentrales semantisches Repositorium, welches die von den Mitgliedstaaten im Rahmen der nationalen Evidence Surveys beigebrachten Informationen enthält.
2. CS Admin Tool
Ein Frontend zur Eingabe der Nachweisdaten, Registerinformationen und Mappings.
3. Air Tables¹⁹
Ergänzende Daten.
4. Evidence Explorer²⁰
Tool zum Anzeigen der produktiven Daten des Semantic Repository.
5. Das Exchange Data Model²¹
Ein auf CCCEV-AP basierende Erweiterung für SDG (siehe TDDs).
6. Zahlreiche Kernvokabulare²² und Taxonomien²³ zur Nachnutzung für strukturierten Datenaustausch.
7. LCM-Interface²⁴,
Eine XML-Schnittstelle, die Daten aus dem Evidence Broker und Dataservice (CS Admin tool) exportieren kann.

Dies entspricht im W3C Semantic Web Stack folgender Abdeckung:

¹⁸ Semantic Repository Beschreibung in den EU TDDs: <https://ec.europa.eu/digital-building-blocks/sites/display/TDD/3.3+-+Semantic+Repository+%28SR%29+-+Snapshot+Q3>

¹⁹ Airtable der EU OOTS subgroup map <https://airtable.com/appXo8kHyglGEFAdj/shrNJuw4qfOLoAtkO/tbljiEble-ZEs4Qlp4?viewControls=on..>

²⁰ Evidence Explorer, Tool der EU KOM zum Anzeigen der im Produktivsystem hinterlegten Daten der Mitgliedstaaten, <https://oots.pages.code.europa.eu/evidence-explorer/ee-app/#/evidences>

²¹ EU OOTS Exchange Data Model (EDM), <https://ec.europa.eu/digital-building-blocks/sites/pages/viewpage.action?pageId=797081666>

²² Kernvokabulare <https://interoperable-europe.ec.europa.eu/collection/semic-support-centre/core-vocabularies>

²³ Weitere Taxonomien und Authority Tables für strukturierten Datenaustausch z.B. beim European Learning Model oder beim zukünftigen Population data model. <https://op.europa.eu/de/web/eu-vocabularies/vocbench>

²⁴ LCM-Interface, XML-Schnittstelle um Daten im CCCEV-AP Format aus dem Evidence Broker zu erhalten, <https://ec.europa.eu/digital-building-blocks/sites/pages/viewpage.action?pageId=706382130>

SDG EU OOTS Komponenten im Semantic Web Stack

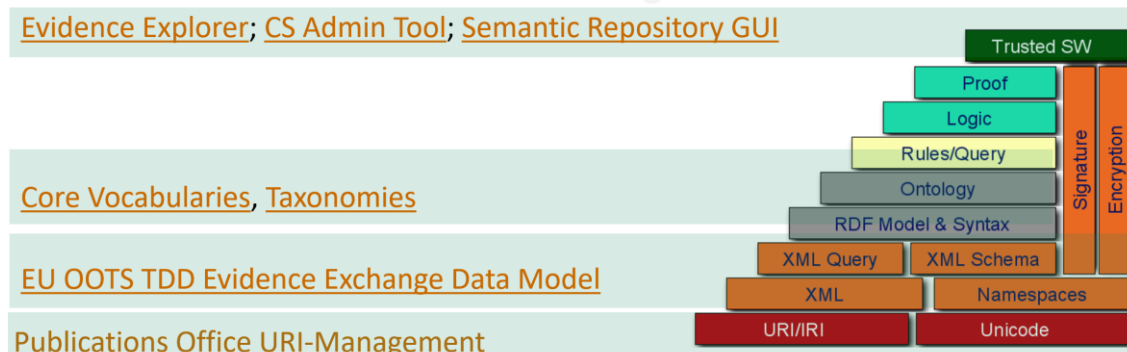


Abbildung 7: EU OOTS im Semantic Web Stack

In der EU-Unterarbeitsgruppe „MAP“ werden die Daten und Mappings erarbeitet und pan-europäisch abgestimmt. Im Wissensgraph des Semantic Repository des EU OOTS werden diese Angaben zu Nachweisen, Providern, Procedural Requirements und Member State procedures dann gehalten und in die verschiedenen Systeme ausgespielt (Evidence Explorer; CS Admin Tool (Evidence Broker), Airtables)

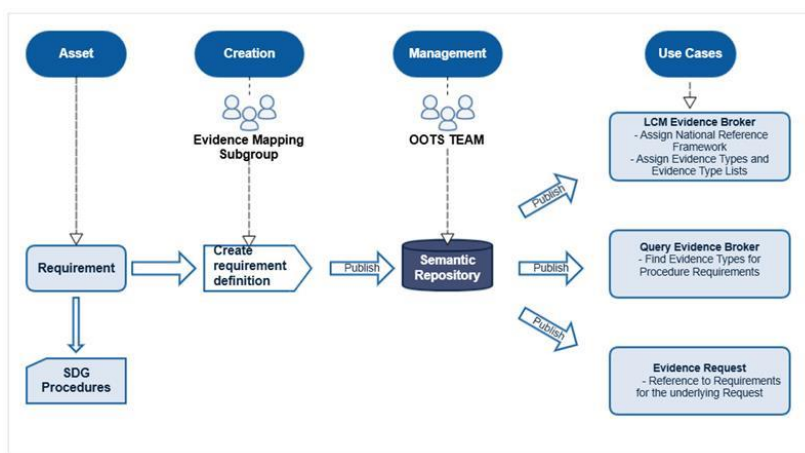


Abbildung 8: Rolle des Semantic Repositories beim EU OOTS

Zur Anknüpfung für RegCheck bietet sich der für das „Semantic Repository“ als einer von drei Use Cases ausgemachter, sog. „Evidence Broker“ an. Aus diesem können mittels LCM-Interface z.B. Daten mit Relevanz für Deutschland aus dem „Semantic Repository“ ausgelesen und in den nationalen Triple Store RegCheck verbracht werden.

ohne EU-Adapter

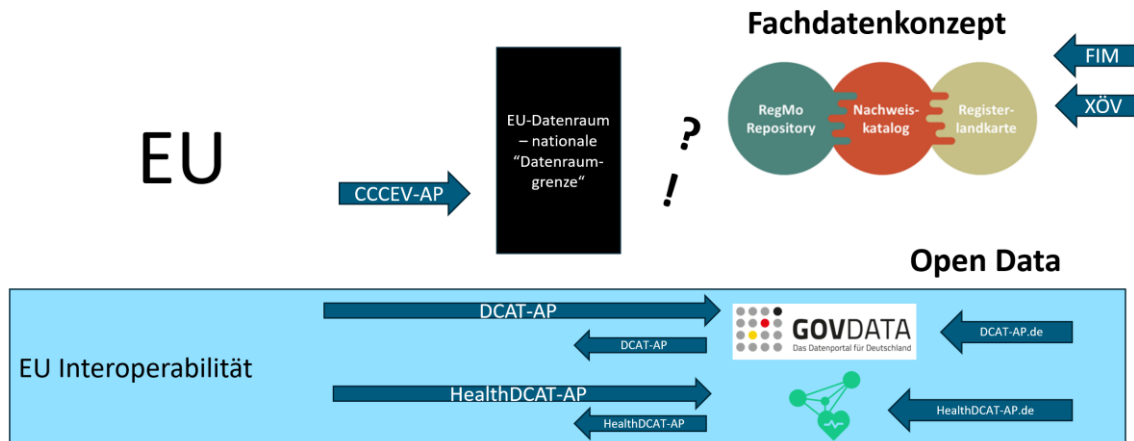


Abbildung 9: EU Datenraum trifft auf nationale Datenraumgrenze

RegCheck könnte hier die Funktion eines „Liason Body“ oder eines „National Access Points“ erfüllen und das semantische Bindeglied zwischen dem EU Evidence Broker und nationaler Ebene sein. Statt wie bisher die EU-Daten manuell für die deutsche Registermodernisierung bereitzustellen, würde man die Nachnutzungsmöglichkeiten von Semantic Web Technologie über die nationale Grenze bringen. Erste bereits im Semantic Web Format agierende nationale Fachportale, wie etwa das Open Data Portal Govdata könnten die EU Evidence Brokerdaten strukturiert und nach FAIR-Prinzipien metadatenattribuiert in der Triple-Auflösung erhalten.

Mit EU-Adapter

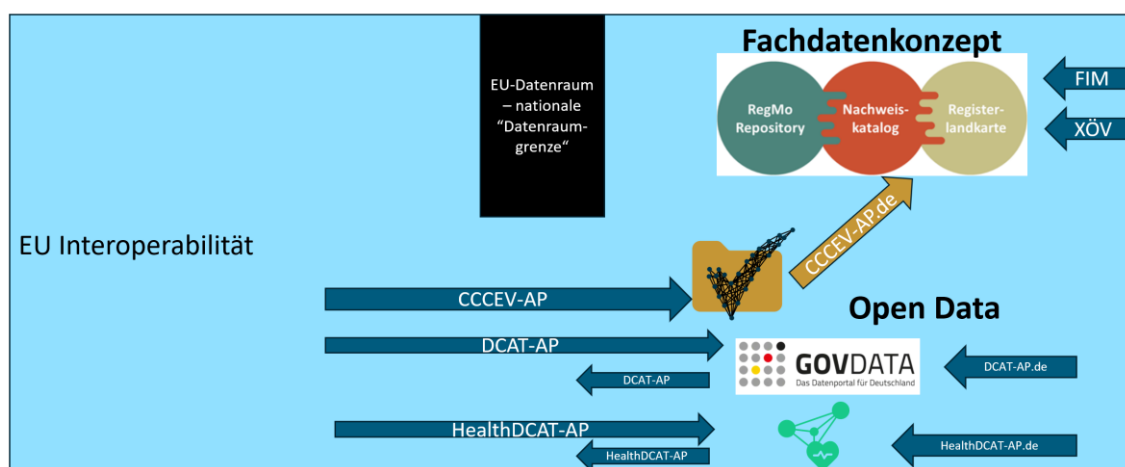


Abbildung 10: EU Datenraum durchbricht dank nationalem CCCEV-AP Profil die Datenraumgrenze

Hier besteht aktuell noch eine konzeptionelle Leerstelle. Die Vorteile wären folgende:

- EU-seitig wird ebenfalls das in RegCheck verwendete CCCEV genutzt; hier könnten also komplexe Datenstrukturen semantisch sinnvoll aufeinander gemappt werden; der Reifegrad D1 auf EU-Seite bindet Core Vokabulare und EU-Taxonomien ein, die auch von der deutschen Ableitung CCCEV-AP.de in ihrer jeweiligen deutschsprachigen Ausprägung genutzt werden können. EU-Interoperabilität qua Design wird möglich.

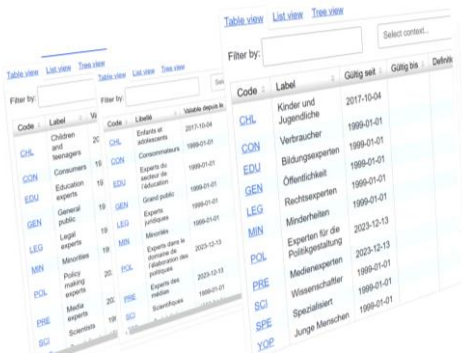


Abbildung 11: Codelisten vom Publications Office als mehrsprachige URI-versehene RDF-Ressourcen

- Komplexität des EU-Modells kann auch im deutschen Föderalismus genutzt werden. Die in der EU beim Evidence Mapping in der Evidence Survey auftretenden Fallkonstellationen sind auch für das NOOTS anzunehmen.

Evidence providers can provide multiple evidence types

Evidence providers may provide many evidence types

Evidence types may be provided by many evidence providers

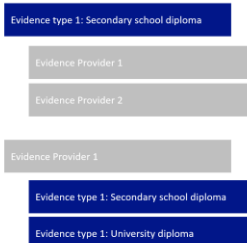


Abbildung 12: beispielhafte Herausforderung von n:n Beziehungen zwischen Providern und Nachweisen im EU OOTS

Nicht zuletzt kann vom informatisch höherwertigen Triple-Format der EU stets in das Fachdatenkonzept mit seiner geringeren semantischen Tiefe konvertiert werden. Damit auf nationaler Ebene die EU-Daten hinzugespeichert werden können, setzte sich das SDG-Team des Programmbereiches EU-OOTS schon zu Beginn der FDK-Konzeption für die Aufnahme von EU-Identifiern ein.

Durch die Nutzung des RegCheck Wissensgraphen mit den dort importierten EU-IDs können die Daten nun weit hinter der Landesgrenze erst in die Welt der klassischen relationalen Datenbanken runterkonvertiert werden. Deutsche SDG-Daten würden somit in Teilen

umgeschlüsselt und auf deutsche Bedürfnisse konvertiert die Registermodernisierung einfließen. Eine fachliche Überforderung des FDK-Prozesses durch ausländische Nachweisinformationen, durch internationale Wertelisten oder FDK-fremde internationale Internetstandards (IANA Mediatype) könnte durch ein dem FDK-vorgelagertes Konvertieren und Umschlüsseln auf die beim FDK verwendete Wertelisten verringert werden. Eine gewisse Abkopplung vom EU-Releasezyklus wäre möglich. So wie zur Laufzeit die IP und die RDN es vermag deutsche Komplexität vor dem EU-OOTS zu kaskadieren, so ist die CCCEV-AP.de die Schnittstelle zur Mapping-Zeit das medienbruchfreie Mittel der Wahl zur Übertragung von Nachweisdaten aus dem EU-Evidence Broker in nachgelagerte nationale Systeme (FDK, OZG-IP, RLK, VIP).

3.7.2.3 Aufbau des Media Type im Nachweisangebotsschema

Im Nachweischema des Data Providers wird auch der verfügbare „Media Type“ (wie in XNachweis deklariert) beschrieben. Der Media Type wird dabei in einem Schlüsselwert abgebildet. Jeder Media Type, den der DP unterstützt, wird auf „1“ gesetzt. Ein Media Type, der nicht unterstützt wird, wird auf „0“ gesetzt. Mindestens ein Media Type muss unterstützt werden. Daraus ergibt sich eine begrenzte Zahl an Kombinationen.

Zu beachten ist das der Media Typ nur auf Nachweis Typen des Reifegrad B und C handelt. Also immer vollständige Dokumente übermittelt werden. Der Reifegrad D1 wird nicht im Media Typ abgebildet.

jpeg	jpg	json	png	pdf	xml	svg	Dez. Wert
0	0	0	0	0	0	0	0
0	0	0	0	0	0	1	1
0	0	0	0	0	1	0	2
0	0	0	0	0	1	1	3
0	0	0	0	1	0	0	4
0	0	0	0	1	0	1	5
0	0	0	0	1	1	0	6
0	0	0	0	1	1	1	7
0	0	0	1	0	0	0	8
0	0	0	1	0	0	1	9

Abbildung 13: Codeliste "Media Type" im FDK 0.4 mit platzsparendem Byte-Kombinationsansatz aus Zeiten des Compilerbaus

- Weitere Datenstrukturen wie Registerlandkarte, Verwaltungsinformationsplattform oder auch Destatis könnten ebenfalls die weitergegebenen Daten nutzen.
- Als positiver Nebeneffekt wären die bisher als Airtables geführten Daten als Open Data nutzbar zu machen (entsprechende Anforderung in der EU Subgroup Map ist bereits erfolgt).

Die Kette der Datenmodelle würden schematisch in folgender Form ineinandergreifen:

Voraussetzung wären hier der in Abbildung 10 dargestellte Konverter zwischen EU-Datenmodell und dem RegCheck-Datenmodell (mit dem Vorteil der geteilten CCCEV-Grundlage) und zwischen dem RegCheck-Datenmodell und nachnutzenden Datenmodellen wie dem Fachdatenkonzept. Erfolgreich wird dieser Durchbruch der nationalen Datenraumgrenzen seit einigen Jahren mit DCAT-AP.de, dem deutschen Metadatenapplikationsprofil für offene Daten der Verwaltung gelebt. Neuere Entwicklung (HealthDCAT-AP.de²⁵) versuchen diesen Erfolg für andere Datenräume (hier: europäischer Gesundheitsdatenraum) zu wiederholen.

Regcheck kann im europäischen Datenraum (Data Spaces) für Daten der öffentlichen Verwaltung zu Registern, Nachweisen und Nachweisbedarfen als ein nationales Registerlandkarten- und Onlineserviceverzeichnis in Ergänzung zum Fachdatenkonzept dienen. Es kann die Heterogenität im Föderalismus, Unterschiede von Gesetzesgrundlagen und gelebter

²⁵ HealthDCAT-AP.de, Spezifikation für den nationalen Gesundheitsdatenraum im Rahmen der PostCovid Data Challenge des BMI und BMWK

Verwaltungspraxis sowie auch ungenaue bzw. unscharfe Mappings abbilden und macht sich so als Vorstufe für spätere Harmonisierungsschritte besonders nützlich.

Nachweisdaten im europäischen Datenraum

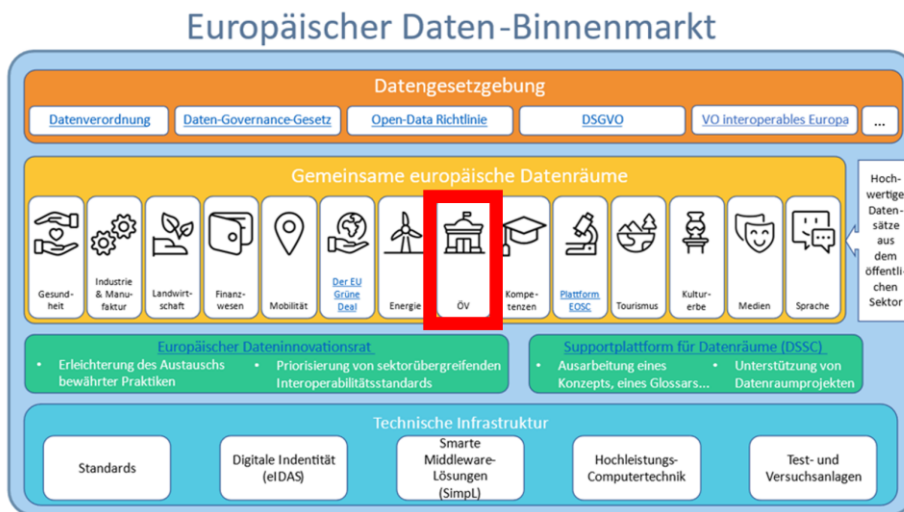


Abbildung 14: Datenräume der EU, darunter "Öffentliche Verwaltung"

3 Pflegekonzept

Das vorliegende Pflegekonzept gibt einen ersten Ausblick darauf, wie Betrieb und Pflege des Wissensgraphen nach Inbetriebnahme geregelt werden können. Der Wissensgraph und das dazugehörige Datenmodell werden vom Betreiber gepflegt und weiterentwickelt. Sie sollen regelmäßig und anlassbezogen in Bezug auf Aktualität geprüft und angepasst werden. Darüber hinaus werden Change Requests eingearbeitet. Welchen organisatorischen Kontexten eine Pflegestelle unterliegt, kann an dieser Stelle nicht geklärt werden. Diese Entscheidung muss zu Beginn der Inbetriebnahme (Verortung von Betrieb und Pflegeinstanz) geklärt werden. Das vorliegende Kapitel nimmt eine generelle Beschreibung der Aspekte, wie die Pflege eines Wissensgraphen mit dem dazugehörigen Datenmodell zu handhaben ist, vor.

3.1 Pflege des Wissensgraphen und des Datenmodells

Die zentrale Pflegestelle ist hauptverantwortlich für die Umsetzung von Anpassungen im Datenmodell und am Wissensgraphen. Sie berücksichtigt Best Practices der Erstellung und Bereitstellung von Datenmodellen, wie etwa die im Deliverable 3.2 Methodenpapier referenzierten SEMIC Style Guides²⁶. Damit einher gehen folgende Aufgaben:

- Die Pflegestelle verantwortet die Aktualisierung hinterlegter Standard-Vokabulare (Bsp: DSMeld) bei Versionsaktualisierungen dieser und sondert veraltete Versionen aus
 - Konzept "veralteter Standard wird noch eine Version mitgeführt und wird dann entfernt"
 - Versionsaktualisierungen werden durch die Pflegestelle regelmäßig geprüft. Dies schließt die Einbindung der Vokabulare, bspw. in Nachweisen und Nachweisanforderungen ein.
 - Bei Anpassungsbedarfen, die sich etwa durch Entfernen verwendeter Elemente mit einem Versionswechsel ergeben, wird ein Change Request erstellt.
 - Veraltete Versionen werden i.d.R. einen Releasezyklus weitergeführt, aber entsprechend markiert (deprecated). Dies ermöglicht eine Umsetzung von Change Requests in geordneten Bahnen.
 - Bei Aktualisierungen der Vokabulare auftretende Anpassungsbedarfe im Wissensgraphen (verwendete Datenfelder in Nachweisen bspw.) werden durch die Pflegestelle aufbereitet:

²⁶ SEMIC Style Guide for Semantic Engineers <https://semiceu.github.io/style-guide/1.0.0/index.html>

- Durch die Pflegestelle werden, sofern fachlich möglich, Beispiellösungen entwickelt, um eine zielgerichtete Bearbeitung des Change Requests zu ermöglichen.
- Theoretisch ist durch den Wissensgraph abbildbar, welche Zusammenhänge durch das Entfernen eines Elements Bearbeitungsbedarfe entwickeln. Perspektivisch sollte dies in der Tool-Lösung berücksichtigt sein.
- Die Pflegestelle sorgt auch unter Beachtung der im URI-Konzept aufzustellenden Regeln für eine Vergabe von weltweit eindeutigen IDs im Wissensgraphen.
- Die Pflegestelle führt in regelmäßigem Turnus Change Advisory Board - Meetings / Expertengremiums-Sitzungen durch. Diese Praxis ist aus dem XÖV-Bereich etabliert und sorgt für eine Abstimmung mit der Fachlichkeit:
 - Besonderheit: Durch die fachübergreifende Aufstellung von RegCheck und auch durch die Aufstellung für verschiedene Ebenen (Kommunal, Landesebene, Bundesebene) setzt sich das Board aus einem größeren Kreis zusammen, für den konkreten Termin werden explizit vor allem Vertreter:innen der Fachlichkeit(en) geladen, die von den Änderungen betroffen sind
 - Ferner können alle Mitglieder dieses Boards Diskussionsbedarfe anmelden

Die Projektgruppe „Semantische Interoperabilität“ des IT-Planungsrates führt im Bereich „Verstetigung von Strukturen“ als Best Practices auf:

Eine zentrale Koordinations-, Anlauf- und Beratungsstelle für Interoperabilität, Standards und Terminologien (Gesetzliche Verpflichtung durch EU) initiiert Abstimmungen und Maßnahmen zur Förderung von Interoperabilität.²⁷

Vor diesem Hintergrund ist auch das Releasemanagement für das Datenmodell zu betrachten. Da strukturelle Änderungen und Anpassungen am Datenmodell planbar vorgenommen werden müssen, erfordert es einen fixierten, regelmäßigen Releasezyklus. Vorgestellt wird hier ein fixer Veröffentlichungstermin in halbjährlichem Turnus, um nutzenden Einrichtungen entsprechende Vorbereitungszeit zur Einbindung zu ermöglichen. Dies wird aktuell erfolgreich von Standards der Innenverwaltung praktiziert. Gesetzliche Änderungen finden Anfang des Jahres und Mitte des Jahres Eingang in die Datenaustausch-Spezifikationen (XMeld, XPersonenstand, XAusländer).

Für die Pflege des Wissensgraphen sind verschiedene Lösungswege vorstellbar, die von der avisierten Nutzungsweise abhängen:

- a) Im Fall eines zentralen Hostings des Wissensgraphen sind feste Releasezyklen nur in Kombination mit Aktualisierungen des zugrunde liegenden Datenmodells von Relevanz. Für diesen Fall sollten Zeiträume eingeplant werden, in denen keine Anpassung

²⁷ Best Practices zu Verstetigung der IT-PLR Projektgruppe zu „semantische Interoperabilität“ https://gitlab.opencode.de/it-planungsrat/pg-sio/konsultationsprozess/-/issues/1#note_199531

des Wissensgraphen durch Nutzende möglich ist. Abseits von diesen wäre der Wissensgraph effektiv in einem kontinuierlichen Update-Prozess, Änderungen durch Nutzende des Wissensgraphen, für die kein manueller Prüfprozess erforderlich wird, könnten so live eingepflegt werden.

- b) Sofern der Wissensgraph in einer lokalen Version in verteilten Rechenzentren für die jeweiligen Einrichtungen vorgehalten wird, ist hingegen ein häufigerer Releasezyklus erforderlich. Es ist Nutzenden in diesem Fall nicht zumutbar, im ungünstigsten Fall sechs Monate auf die Einarbeitung von Anpassungen zu warten – eine solche Verzögerung würde rasch die Akzeptanz der Anwendung senken, selbst wenn die Änderungen in der lokalen Fassung bereits nutzbar wären. Hintergrund ist, dass auch bei lokaler Nutzbarkeit eine Interoperabilität der Formulare im Grunde nicht möglich wäre.

Eine zentrale Pflege der im Wissensgraph verwendeten Taxonomien unterstützt eine zeitnahe Bereitstellung von durch Verwaltungsmitarbeitende initiierte Änderungen.

Wird der Wissensgraph hingegen lokal vorgehalten brauchen wir sogar häufigere Release Zyklen als halbjährlich, denn wenn da z.B. ein Sachbearbeiter ein neues Formular für einträgt, will er/sie ja dann nicht ein halbes Jahr warten, bis das auch benutzt werden kann.

Im Gegensatz dazu wird auf Ebene der Instanzdaten kein fester Zyklus avisiert. Hier handelt es sich um einen fließenden Prozess, in dem Informationen tagesaktuell eingearbeitet werden. Der Wissensgraph (Struktur + Daten) ist auf Github frei verfügbar²⁸.

3.2 Einarbeitung Change Requests

Die Pflegestelle begleitet außerdem die Einarbeitung von Change Requests.

- Change Requests können generell durch alle Stakeholder eingereicht werden
 - Betrifft: Fachvertreter, aber auch und insbesondere Verwaltungen, die mit dem Wissensgraphen arbeiten und diesen erweitern wollen
- Ergänzungen des Wissensgraphen sollen in der Regel durch die Verwaltung selbst eingereicht werden können. Die Pflegestelle prüft ausschließlich bei neu hinzugefügten oder fachfremd eingefügten Datenfeldern
 - Z.B.: Wird aus dem Bereich des Meldewesens ein neues Formular angelegt, welches durchgängig die bereits aus DSMeld enthaltenen Datenfelder enthält, bedarf dieses Formular keiner weiteren Prüfung. Würde ein Datenfeld aus einer anderen Fachlichkeit, z.B. aus dem Finanzwesen, hinzugefügt, wird dies geprüft.

Für die Nutzung von Datenfeldern in Formularen gibt es drei vorstellbare Szenarien:

²⁸ Wissensgraph der deutschen Verwaltung; hier: Registermodernisierung:

https://github.com/RegCheck2024/RegCheck/blob/main/rdf/RegCheck2024_Graph_20241126.ttl

- a) Das Formular nutzt ausschließlich bereits vorhandene Datenfelder aus der zur jeweiligen Verwaltungseinrichtung gehörigen Fachlichkeit. In diesem Fall ist bei einer Anpassung keine manuelle Prüfung durch die Pflegestelle erforderlich.
- b) Das Formular nutzt ausschließlich bereits vorhandene Datenfelder, integriert jedoch auch Felder aus anderen Fachlichkeiten. In diesem Fall wird das Change Request durch die Pflegestelle geprüft und bei Bedarf eine Rückfrage gestellt. In der Regel kann davon ausgegangen werden, dass dieser Prozess kurzfristig und bilateral abläuft und anschließend eine Freigabe erfolgt. In speziellen Fällen kann eine Klärung im Change Advisory Board erforderlich sein.
- c) Das Formular nutzt neue Datenfelder, die bislang noch nicht im Wissensgraph enthalten sind. In diesem Fall prüft die Pflegestelle zunächst intern auf ggf. vergleichbare Datenfelder aus verschiedenen Vokabularen. Das Change Request wird im Rahmen des Change Advisory Boards diskutiert und ggf. integriert. Vorgeschlagene Taxonomien, Vokabulare und Applikationsprofile werden vor der Nachnutzung einem CAMSS-Assessment unterzogen. Siehe auch das in der Umfeldanalyse referenzierte CAMSS-Assessment von CCCEV²⁹.

*** A40 - Is the specification included in an open repository/catalogue of standards at national level?**

EIF Recommendation 23: Consult relevant catalogues of standards, specifications, and guidelines at the national and EU level, in accordance with your NIF and relevant DIFs, when procuring and developing ICT solutions.

EIF Recommendation 6: Reuse and share solutions, and cooperate in the development of joint solutions when implementing European public services.

- ☐ Not Answered
- ☐ Not Applicable
- ☐ NO
- ☒ YES

Abbildung 15: Prüfkatalogs CAMSS zum Prüfen der Governance und anderer Aspekte vorgeschlagener neuer Vokabulare

3.3 Management der Pflege

3.3.1 Regelbetrieb, Support und Beratung

Der Betreiber kommuniziert die Änderungen am Modell und am Wissensgraphen öffentlich und stellt diese über eine Informationsplattform (Website) zur Verfügung. Dort ist es auch möglich, Fragen zu stellen, Hilfe anzufordern oder Änderungen vorzuschlagen.

Der Betreiber übernimmt ebenso das Konfigurationsmanagement, indem er

- die releasebezogenen Konfigurationen dokumentiert und verwaltet

²⁹ CAMSS Assessment von CCCEV 2.0.0.0 https://interoperable-europe.ec.europa.eu/sites/default/files/distribution/access_url/2023-05/ddeoce8c-855d-4cd8-b8f6-006cf57aff65/CAMSS_Assessment_CCCEV{EIFScenario_v1.0.0.pdf

- sowie den Status der Konfigurationselemente überwacht

3.3.2 Stakeholdermanagement

Im Regelbetrieb ist es wichtig, die Bedürfnisse und Erwartungen der beteiligten Parteien (Stakeholder) zu berücksichtigen. Durch das Stakeholdermanagement sollen die Interessen der Stakeholder mit den Zielen des Betriebs des IT-Standards in Einklang gebracht werden. Dabei wird eine stabile Kommunikation zwischen dem Change Advisory Board, als umsetzende Einheit der Änderungen, und den Stakeholdern etabliert. Denn in der Kommunikation ist es essentiell, die Anliegen der Stakeholder durchgängig zu adressieren und sie in die Entscheidungsfindung einzubeziehen.

Im Vorfeld ist ein Überblick über die relevanten Personen und Gruppen herzustellen. Diese können nach Interesse und Einfluss auf den Betrieb des IT-Standards gruppiert werden, um so durchgängig und angepasst mit den Stakeholdern zu kommunizieren.

Es wird empfohlen, dass Change Advisory Board sowohl mit technischen Mitarbeitern als auch Entscheidungsträgern zusammen zu setzen. Wenn nötig können Fachexperten hinzugezogen werden, um detaillierte Änderungsanfragen zu bewerten.

3.3.3 Änderungsmanagement

Das Änderungsmanagement ist dafür verantwortlich, Änderungswünsche an dem IT-Standard geordnet und effektiv zu bearbeiten. Es sorgt dafür, dass alle Beteiligten transparent und passend zur Anforderung zusammenwirken.

Es wird ein **Releaseplan** erstellt, aus dem hervorgeht, welche Änderungen in welchem Release enthalten sind. Der Plan regelt folgende Punkte:

- Welche Änderungen werden zu welchem Zeitpunkt umgesetzt
- Festlegung Release-Zyklus
- Release-Nummer
- Release-Datum
- Liste umzusetzender Änderungen
- Freigabe der Releases
- Release-Zeitpunkt

Hierfür stellt der Betreiber eine Plattform zur Verfügung und organisiert so ein öffentlich online zugängliches und einsehbares Änderungsmanagement. Ziel ist es, eine umfassende Transparenz herzustellen, in der angefragte, geplante, in Bearbeitung befindliche und abgeschlossene Änderungen einsehbar sind.

4 Designentscheidungen in der Modellierung

Wann immer ein Modell als ein abstrahiertes Modell der Realität erstellt wird, müssen eine Reihe von Abwägungen und Entscheidungen getroffen werden. Gründe hierfür sind unter anderem die Komplexität der Realität - die nie in Gänze abgebildet werden kann -, das Ziel des Modelles, die Verfügbarkeit von Daten sowie die Zielgruppenorientierung.

Auch der im Projekt RegCheck entstandene Wissensgraph bildet hier keine Ausnahme. Die beiden EU-Vokabulare CPSV und CCCEV, der DSMeld, FIM, die Verwaltungsdateninformationsplattform, die Registerlandkarte, die Rechtsgrundlagen sowie zahlreiche Eigenschaften der deutschen Verwaltung wurden allesamt in dem Wissensgraphen abgebildet. Bei diesem Prozess ergaben sich viele Fragestellungen, die thematisieren, wie etwas am akkuratesten im Wissensgraphen modelliert werden kann. Außerdem musste abgewogen werden, wie am pragmatischsten vorgegangen wird, sodass der Pflegeaufwand minimiert und der Mehrwert für die Anwendung sowie die Zielgruppenorientierung maximiert werden. Die Entscheidungen, die im Rahmen der Modellierung getroffen wurden, wurden zur Nachvollziehbarkeit dokumentiert und sind in diesem Dokument zu finden.

Zusätzlich wurden auch bisher offene Thematiken hier transparent dargestellt und sollten als eine erste Aufgabenliste für empfehlenswerte Erweiterungen und Ergänzungen des Wissensgraphen dienen. Auch wurden im Rahmen des Projektes Erkenntnisse gewonnen, welche auch für die EU-Vokabulare CCCEV und CPSV relevant sind, sodass diese in Form von Issues an die SEMIC Community weitergeleitet wurden. Diese sind ebenso in diesem Abschnitt enthalten.

4.1 Ontologie und Modellstruktur

Titel	Explizite Modellierung von „hasSupportingEvidence“ und „supportsRequirement“
Beschreibung	<p>Es stellt sich die Frage, ob eine explizite Modellierung von „hasSupportingEvidence“ und „supportsRequirement“ benötigt wird.</p> <p>Im IST-Zustand des Modells werden beidseitige Referenzierungen von je Information Requirement und Evidence auf Information Concept erstellt. Eine 1:1 Beziehung im Sinne von „dieses Evidence bezieht sich auf dieses Information Requirement“ ist vermutlich nicht mehrwertstiftend und verursacht einen erheblichen Pflegeaufwand.</p>
Lösungsvorschläge	<ol style="list-style-type: none"> 1) Keine 1:1 Beziehungen zwischen Evidence und Information Requirement modellieren. 2) 1:1 Beziehungen zwischen Evidence und Information Requirement modellieren.
Darstellung der Lösung	<p>Entscheidung für Lösungsalternative 1). Beide Kanten sind nun im Graphen optional. Dies ist für die spätere Pflege und Inbetriebnahme des Graphen von Vorteil, da so weniger verpflichtender Aufwand gegeben ist. Zugleich kann es bei Bedarf hinzugefügt werden.</p> <p>CCCEV folgt einer Logik auf Dokumentenebene und eine solche Darstellung einer 1:1 Beziehung ist nicht sinnvoll.</p> <p>Die Beziehung zwischen Information Requirement und Evidence, also das eine Informationsanforderung im Verwaltungsprozess etwas aus einem Nachweis benötigt, ist auf Datenfeldebene über die Information Concept gewährleistet.</p>

Titel	Betroffene Personen in DSMeld / im Verwaltungsprozess
Beschreibung	<p>Hier ist die Herausforderung zu entscheiden, wie mit der Personen-Logik von DSMeld umgegangen wird. In DSMeld werden neben der meldepflichtigen Person auch Daten zu weiteren Personen genannt wie z.B. Vorname des Kindes, Geburtsdatum des Ehepartners. Jedoch sind die Datenkränze der weiteren Personen deutlich kleiner als die Datenkränze der meldepflichtigen Person und besitzen z.B. keine Angaben zur Adresse.</p> <p>Zugleich ist hiermit auch die von DSMeld losgelöste Frage verbunden, wie mit Daten Anforderungen in einem Prozess in Bezug auf Personen, die nicht</p>

	der Hauptakteur des Prozesses sind, umgegangen werden soll und wie dies modelliert werden kann.
Lösungsvorschläge	Es kann das DSMeld Blatt 001 „Betroffene Person“ verwendet werden. Dieses bietet eine Auswahl zwischen verschiedenen Personen: z.B. meldepflichtige Person, gesetzlicher Vertreter – Vater, Elternteil. Dieses DSMeld Blatt kann z.B. mit den Adressangaben kombiniert werden, um die Adresse eines Kindes auszudrücken.
Darstellung der Lösung	Sofern Datenfelder (z.B. Vorname des Kindes) vorhanden sind, wird auf diese referenziert. Für Datenfelder ohne einen speziellen Personenbezug werden Konzeptlisten erstellt, welche Personen des DSMeld Blattes „Betroffene Person“ enthalten. Sofern ein „Information Requirement“ Daten einer anderen Person als der Hauptperson des Prozesses benötigt, wird dies mit einem „betrifftAbweichendePerson“ als Verweis auf die konkrete abweichende Personengruppe aus der Konzeptliste, wie z.B. Lebenspartner, dargestellt. Aus dem so gekennzeichneten Information Requirement wird dann sowohl auf die personenspezifischen Felder (z.B. Vorname des Kindes) als auch auf die nicht personenspezifischen Felder (also die allgemeinen Felder für die Adresse) verwiesen.

Titel	Umgang mit verschiedenen IDs
Beschreibung	Neben DSMeld berücksichtigt das Modell auch weitere Vorhaben wie die Verwaltungsinformationsplattform, die Registerlandkarte und FIM. All diese Vorhaben besitzen eigene IDs für die Datenfelder. Es ist zu eruieren, wie mit den verschiedenen IDs im Modell umgegangen wird.
Lösungsvorschläge	Siehe Lösung
Darstellung der Lösung	<p>Es wurde entschieden folgende IDs zu berücksichtigen:</p> <ol style="list-style-type: none"> 1. FIM – da es eines der großen Standardisierungsvorhaben der deutschen Verwaltung darstellt; 2. DSMeld – da es als standardisierte Grundlage für die erste Datenmodellierung und die entsprechenden Datenfelder diene; 3. VIP – da es eine umfassende Auflistung der deutschen Register darstellt; 4. RLK – da es als wichtiges Projektmanagement- und Dokumentationswerkzeug für die Registermodernisierung dient <p>Weitere IDs können bei Bedarf ebenso hinzugefügt werden.</p> <p>Die IDs wurden als blank nodes an das jeweilige Information Concept „angehängt“. Hierzu wurde das Asset Description Metadata Schema (ADMS) verwendet.</p>

Titel	Umgang mit Hierarchie der Gebietskörperschaften
Beschreibung	Es wurde ein Schlüssel gesucht, welcher die deutsche föderale Struktur abdeckt. Aktueller Fokus ist nur auf deutsche föderale Strukturen. Föderale Strukturen anderer EU-Länder werden nicht berücksichtigt, da zunächst eine nationale Lösung entwickelt wird. Hier stehen verschiedene Alternativen zur Auswahl
Lösungsvorschläge	<p>Als Folge davon wird aktuell ein anderer Schlüssel gesucht, welcher die deutsche föderale Struktur abdeckt. Aktueller Fokus ist nur auf deutsche föderale Strukturen. Föderale Strukturen anderer EU-Länder werden nicht berücksichtigt, da wir eine nationale Lösung entwickeln. Hier stehen verschiedene Alternativen zur Auswahl.</p> <ol style="list-style-type: none"> 1. NUTS ist eine europäische Klassifikation, die Gebietseinheiten in verschiedene, hierarchische Stufen einteilt u.a. nach Verwaltungsebenen, aber auch nach Bevölkerungsgrößen; 2. Die von DCAT-AP.de genutzte Einteilung der deutschen Verwaltungseinheiten.
Darstellung der Lösung	<p>Es wird der DCAT-AP Schlüssel verwertet, da das Modell eine deutsche Lösung darstellt und somit ausreichend ist.</p> <p>Zugleich kann NUTS nicht verwendet werden, da aus Sicht der registerführenden Stellen ein Register eventuell in zwei verschiedenen Gebieten nach NUTs liegen kann, je nachdem wie bevölkerungsreich die Kommune/Stadt ist und ihrer dementsprechenden Einteilung in NUTS.</p>

Titel	Schreibweise des Namens eines Nachweises
Beschreibung	Es wird betrachtet, wie genau die Nachweise aus dem Melderegister in dem Modell benannt werden sollen.
Lösungsvorschläge	Abgewogen wurde, ob das Bundesmeldegesetz im Namen der Nachweise ausgeschrieben oder als „BMG“ abgekürzt verwendet wird.
Darstellung der Lösung	Das Bundesmeldegesetz im Namen der Nachweise wird abgekürzt verwendet, da so der Name kürzer ist und dennoch Eindeutigkeit besteht. Ein Beispiel für ein Nachweisname im Modell ist "Einfache Melderegisterauskunft nach § 44 I BMG".

Titel	„Publisher“ von „Evidences“
Beschreibung	CCCEV beschreibt die Klasse „Publisher“ als „legal authority responsible for the evidence“, im Modell wird sie als dcterms:publisher dargestellt. Im Falle

	des Meldewesens ist zu prüfen, ob für die Nachweise aus dem Melderegister hier Melderegister oder Einwohnermeldeamt als „publisher“ eingetragen werden.
Lösungsvorschläge	<ol style="list-style-type: none"> 1. Melderegister wird als „Publisher“ eingetragen 2. Einwohnermeldeamt wird als „Publisher“ eingetragen
Darstellung der Lösung	In dem beschriebenen Fall wird Lösung 2 verwendet, das Einwohnermeldeamt als „Publisher“ eingetragen, da dies die Behörde ist, die für den Nachweis verantwortlich ist und somit der Beschreibung aus CCCEV entspricht.

Titel	Angabe von Nutzungsgründen
Beschreibung	Durch die Ergebnisse der Interviews kann im Modell angegeben werden, welches Datenfeld für welche Verwaltungsleistung aus welchem Grund verwendet wird. Um dies auszudrücken, kann eine entsprechende Eigenschaft „Nutzungsgründe“ eingeführt werden. Offen ist jedoch, wo diese Klasse im Modell verortet wird. Grundsätzlich ist eine Repräsentation nötig, bei der die Aussage „Folgende Informationsanforderung benötigt folgendes Datenfeld ausfolgendem Grund.“ möglich ist, die Aussage bezieht sich somit auf die Verbindung zwischen einem spezifischen Information Requirement und einem universellen Information Concept.
Lösungsvorschläge	<p>Die Lösung beinhaltet die Nutzung von blank nodes. Diese können prinzipiell an zwei Stellen eingesetzt werden:</p> <ol style="list-style-type: none"> 1. Die Eigenschaft Nutzungsgründe kann auf Ebene des „Information Concept“ modelliert werden; 2. Die Eigenschaft Nutzungsgründe kann auf Ebene des „Information Requirement“ modelliert werden.
Darstellung der Lösung	<p>Die Information über Nutzungsgründe für ein bestimmtes Datenfeld wurde auf Ebene der „Information Concepts“ hinterlegt. Die Nutzungsgründe sind quasi Attribute der universellen Information Concepts, die jedoch auf die jeweilige Verbindung zu einem spezifischen Information Requirement rückverweisen.</p> <p>Eine übergreifende Taxonomie von Nutzungsgründen ist zum aktuellen Zeitpunkt nicht umgesetzt worden. Hier bieten die Ergebnisse der Interviews lediglich einen ersten Einstieg in eine umfassendere Systematik, bei der die Frage der Abstraktion und Systematisierung zu vertiefen ist. Aktuell sind die Nutzungsgründe ohne semantische Struktur als Strings aufgesetzt worden.</p>

Titel	Einheitliche Sprache
--------------	----------------------

Beschreibung	In einem Zwischenstand der Modellierung sind Klassen/Elemente/Attribute teilweise auf Deutsch und teilweise auf Englisch bezeichnet.
Lösungsvorschläge	<ol style="list-style-type: none"> 1. Alle Elemente auf Englisch 2. Alle Elemente auf Deutsch
Darstellung der Lösung	Aktuell besteht weiterhin eine Mischung von englischsprachigen Bezeichnungen aus den EU Core Vocabularies und den deutschen Spezifika.

Titel	Differenzierung zwischen Absatz 1 und 2 von §18 BMG
Beschreibung	In §18 BMG werden in Absatz 1 und Absatz 2 unterschiedliche Datenkränze beschrieben. In Absatz 1 werden die Grunddaten definiert und in Absatz 2 mögliche zusätzliche Daten. Daraus entsteht die Fragestellung, wie dies im Wissensgraphen abgebildet wird bzw. auch wie die/ der Nachweis/e benannt werden.
Lösungsvorschläge	<ol style="list-style-type: none"> 1. Es wird eine gemeinsame „Evidence“ modelliert mit allen Datenfeldern aus Absatz 1 und 2; 2. Es werden zwei „Evidences“ modelliert. Eine „Evidence“ mit den Datenfeldern aus Absatz 1, eine mit den Daten aus Absatz 1 und 2.
Darstellung der Lösung	Es wird ein „Evidence“ mit den Datenfeldern aus Absatz 1 und 2 modelliert und in der Benennung allgemein auf §18 BMG verwiesen. Dieser beinhaltet sowohl Absatz 1 als auch Absatz 2 des §18 BMG. Dies verringert den Modellierungsaufwand, da so nicht 2 verschiedene „Evidences“ als Quelle für Datenfelder angegeben werden müssen, die in Absatz 1 beschrieben sind. Damit wird der Logik gefolgt, dass es sich hier real um einen Nachweis, die Meldebescheinigung, handelt.

Thema	Behörden als Teile einer Behörde sind.
Beschreibung	Es ist unklar, ob im föderalen System der BRD die Behördenstruktur einheitlich modelliert werden kann, als ob z.B. die Behörde, die für Namensänderungen zuständig ist, immer das Bürgeramt ist.
Lösungsvorschläge	Siehe Lösung
Darstellung der Lösung	Es wird die Kante capde:geführtVon“ genutzt, um bei den Registern anzugeben, dass sie von einer „Public Organisation“ geführt werden. Für nachweis-anfordernde Stellen ist eine Abbildung der Hierarchie zwischen nachweis-anfordernden Behörden zum jetzigen Zeitpunkt nicht vorgesehen.

4.2 DSMeld

Titel	Abhängigkeitslogik der DSMeld Blätter untereinander
Beschreibung	Auf den DSMeld-Blättern ist stets vermerkt, wo Abhängigkeiten oder Referenzen zu anderen Blättern bestehen.
Lösungsvorschläge	Siehe Lösung
Darstellung der Lösung	Die Herausforderung stellt sich rein auf der beschreibenden Ebene, nicht jedoch auf der Ebene der Datenmodellierung. Da notwendige Blätter in „Information Concepts“ abgebildet werden, ist es irrelevant, dass diese Referenzen auf andere Blätter beinhalten. Es liegt also kein Informationsverlust vor.

Titel	Umgang mit „Datenkränzen“
Beschreibung	In DSMeld existieren Datenkränze: Zum Beispiel ist die Anschrift ein Datenkranz, der wiederum mehrere Datenfelder (Straße, Hausnummer etc.) umfasst. Es ist zu prüfen, wie diese Hierarchie in dem Modell abgebildet wird.
Lösungsvorschläge	Siehe Lösung.
Darstellung der Lösung	Die Datenkränze werden im Modell berücksichtigt. Die Datenfelder, die Teil eines Datenkranzes sind, werden als „conceptOf“ des Datenkranzes modelliert. Zugleich erfüllt weder diese Verbindung noch die Datenkranz selbst zum jetzigen Zeitpunkt eine funktionale Rolle innerhalb des Modells. Die Logik des Modells bleibt allein auf Ebene der einzelnen Datenfelder.

Titel	Umgang mit BMG/ BMeldDigiV
Beschreibung	Bei einigen Nachweisen könnte als Rechtsgrundlage sowohl auf das Bundesmeldegesetz wie auch auf die Bundesmeldedigitalverordnung verwiesen werden.
Lösungsvorschläge	<ol style="list-style-type: none"> 1. Man verweist allein auf das BMG. 2. Man verweist, wo möglich, auf die BMeldDigiV und für verbleibende Nachweise auf das BMG. 3. Man legt Nachweise doppelt an.
Darstellung der Lösung	Es wurde festgestellt, dass die BMeldDigiV in den hier behandelten Kontexten auf das BMG §18 Abs. 2 referenziert. Aktuell ist dieser Nachweis in Bezug auf die BMeldDigiV erstellt.

Titel	Abgleich mit DSMeld XML
Beschreibung	Neben der PDF-Datei wird zu DSMeld auch eine XML-Datei herausgegeben. Es gilt zu prüfen, ob diese zusätzliche und relevante Informationen enthält.
Lösungsvorschläge	Siehe Lösung
Darstellung der Lösung	Die DSMeld.xml ist die XML-Repräsentation der PDF-Datei. Als zusätzliche, relevante Information sind in der XML für einige DSMeld-Felder die entsprechenden FIM-BOB-Datenfelder hinterlegt. Diese Erkenntnisse lagen bereits vor. Zugleich ist diese Erkenntnis von zusätzlichen Informationen in den XML-Dateien relevant für die Einbindung zukünftiger Fachstandards und kann separate Recherchen ersparen.

Titel	DSMeld1200 "Adresse unbekannt"
Beschreibung	Bei der Meldebestätigung nach §24 BMG wird das Datum „Anschrift – unbekannt“ gemäß DSMeld-Blatt 1200 nicht verwendet, bei anderen Nachweisen aus dem Melderegister wie z.B. der einfachen Melderegisterauskunft jedoch schon. Es muss validiert werden, ob dies korrekt ist.
Lösungsvorschläge	Siehe Lösung
Darstellung der Lösung	Die Lösung lässt sich fachlich herleiten: Die Meldebestätigung hat den Zweck, eine bestimmte Adresse zu bestätigen – daher ergibt die Angabe der unbekannten Adresse in diesem Kontext keinen Sinn. Bei anderen Nachweisen wie der Melderegisterauskunft geht es vielmehr darum, die Information zu erhalten, die zu einer Person im angefragten Melderegister gespeichert sind, daher ist hier die unbekannte Adresse ein sinnvolles Datum.

4.3 FIM

Titel	FIM als (unzureichende) Herleitung für „Information Requirements“
Beschreibung	Aus den FIM-Leistungsbeschreibungen wird nicht erkenntlich, welche Datenfeldern aus den Leistungen für welche Verwaltungsleistungen benötigt werden.
Lösungsvorschläge	Siehe Lösung

Darstellung der Lösung	Bis alle Interviews durchgeführt und ausgewertet worden sind, wurde unter der Prämisse modelliert, dass der gesamte Inhalt der in den FIM-Leistungsbeschreibungen unter „Erforderliche Unterlagen“ aufgeführten Nachweise aus dem Meldewesen benötigt wird. Nach Abschluss der Interviews wurden diese Informationen im Modell berücksichtigt.
-------------------------------	--

Titel	Lange FIM-Beschreibung
Beschreibung	Die Leistungsbezeichnungen I bzw. Leistungsbezeichnungen II der FIM-Leistungsbeschreibungen sind teilweise sehr lang: z.B. „Aufenthaltserlaubnis aus familiären Gründen - Erteilung für den Familiennachzug eines ausländischen Elternteils zu einem minderjährigen Sohn“
Lösungsvorschläge	Siehe Lösung
Darstellung der Lösung	Anstelle der Leistungsbezeichnungen wird die FIM-ID für Public Services verwendet.

Titel	Unzureichend formulierte Nachweisanforderungen in den FIM-Leistungsbeschreibungen
Beschreibung	Oftmals sind in dem Abschnitt „Erforderliche Unterlagen“ die beizubringenden Nachweise nicht konkret genug beschrieben. Beispiele sind hier „oder“-Formulierungen wie „Meldebestätigung oder Melderegisterauskunft“ sowie unpräzise Nachweisanforderungen wie „Nachweis über das Einkommen“, „Nachweis über die allgemeine Lebenssituation“.
Lösungsvorschläge	Siehe Lösung
Darstellung der Lösung	Exaktere Informationen können mit der zuständigen Fachlichkeit erhoben werden im Sinne der Interviews. Der Sachverhalt zeigt die grundsätzlichen Limitationen von FIM-Leistungsbeschreibungen und den Mehrwert von Erhebungen auf Datenfeld-Ebene.

Titel	Umgang mit mehreren Rechtsgrundlagen für eine Leistung
Beschreibung	In dem Bereich „Handlungsgrundlagen“ der FIM-Leistungsbeschreibungen, kann mehr als eine Rechtsgrundlage genannt werden. Dies gilt es im Graphen abzubilden.
Lösungsvorschläge	Siehe Lösung
Darstellung der Lösung	Sofern mehr als eine Rechtsgrundlage für einen Public Service vorliegt, werden diese separat voneinander modelliert.

Titel	Umgang mit „description“
Beschreibung	Die Public Services besitzen aus CPSV heraus ein Element „description“. Es ist zu evaluieren, welcher Teil der FIM-Leistungsbeschreibung hier verwendet werden soll.
Lösungsvorschläge	Siehe Lösung
Darstellung der Lösung	Da die gesamte Leistungsbeschreibung zu umfangreich ist, wird der kürzere Teaser der FIM-Leistungsbeschreibung verwendet, der die Leistung kurz und prägnant zusammenfasst.

4.4 Offene Fragestellungen

Obwohl eine Vielzahl an Lösungen erarbeitet wurden, zeigte sich mit jeder neu hinzugekommenen Fachlichkeit und im Austausch mit den Fachexperten, dass eine Reihe von Themen aktuell noch nicht final gelöst werden konnte. Auch Abhängigkeiten von CCCEV/CPSV sind hierfür ein Grund. Teilweise wurden vorläufig eigene Lösungen entwickelt, die über CCCEV/CPSV hinausgehen. Für Fragen, die nicht allein aus der Schaffung eines deutschen Anwendungsprofils resultieren, sondern grundlegenderen Charakter haben, wurden Änderungsanforderungen an die Editoren der CCCEV/CPSV-Datenmodelle gestellt.

Im Rahmen des Projekts wurde im Rahmen der definierten Ziele und eines begrenzten zeitlichen Rahmens das grundsätzliche Funktionieren eines Datenmodells auf Basis eines Wissensgraphen belegt. Dabei war es unvermeidlich, dass einige Themen zunächst unberücksichtigt bleiben mussten. Diese können jedoch im Rahmen eines möglichen Folgeprojekts vertieft werden. Zudem ist eine kontinuierliche Erweiterung des Modells sinnvoll, um weitere fachliche Perspektiven und Inhalte zu integrieren und seine Anwendbarkeit und den potenziellen Nutzen noch zu vergrößern.

Titel	Korrekte Modellierung von Anträgen im Verwaltungsverfahren
Beschreibung	Es ist unklar, wie mit den europäischen Vokabularien Anträge der deutschen Verwaltung abgebildet werden. Das in CPSV verwendete „Life Event“ entspricht in seiner Definition nicht der Bedeutung eines Antrags. Daher bieten die europäischen Vokabularien aktuell keine adäquate Lösung, um den Antrag im Sinne der deutschen Verwaltung abzubilden.
Lösungsvorschläge	Siehe Lösung

Darstellung der Lösung	<p>Keine zum jetzigen Zeitpunkt. Zugleich ist diese Frage sowohl für das RegCheck Modell als auch für CCCEV relevant.</p> <p>Daher wurden entsprechende Tickets bei CCCEV gestellt. Sofern keine entsprechenden Erweiterungen dort erfolgen, wird eine eigene Erweiterung des RegCheck Modells in Zukunft notwendig sein.</p> <p>Über die Verortung der Klasse „Antrag“ im Datenmodell und die entsprechenden Rückwirkungen auf das bestehende Datenmodell wurden erste konzeptionelle Vorüberlegungen angestellt.</p>
-------------------------------	--

Titel	Unterschiedliche Anforderungen von Verwaltungsverfahren
Beschreibung	<p>Die unterschiedlichen Anforderungen an benötigten Datenfelder sind abhängig von der Person, die das Verfahren einleitet (z.B. ob sie in einer Bedarfsgemeinschaft lebt oder nicht; Kinder hat u.ä.) sowie von konkurrierender Gesetzgebung (z.B. divergierende Gesetzgebung in den Bundesländern).</p> <p>Grundsätzlich geht es darum eine „if“-Beziehung im Modell zu modellieren: Wenn bestimmte Voraussetzungen zutreffen, hat dies Folgen für die Anforderungen an Datenfelder.</p>
Lösungsvorschläge	Siehe Lösung
Darstellung der Lösung	<p>Es konnte ermittelt werden, dass diese Herausforderung das Modell derzeit nicht betrifft. Es existieren keine Abhängigkeiten zwischen Information Requirement und der Person, um welche es im Prozess geht. Nichtsdestotrotz ist diese Frage relevant, sobald Anträge und dort enthaltene Logiken zum Modell hinzugefügt werden. Somit wird diese Frage in potenziellen Erweiterungen des Modells relevant werden.</p> <p>Die Entwicklungen im Kontext des openDVA-Projektes mit „Rule-Mapping“ und allgemein der Ansatz, derartige Fragen aus normenanalytischer Perspektive zu klären, wurden zur Kenntnis genommen. Inwiefern diese für die vorliegende Art der Datenmodellierung nutzbar sind, bleibt offen.</p>

Titel	Level of Authentication bei Evidence
Beschreibung	In CCCEV ist in der Klasse „Evidence“ auch die optionale Angabe zur Vertraulichkeitsstufe möglich. Fraglich ist nun, ob diese Angabe mit in das Modell aufgenommen wird und ob die Vertraulichkeitsstufe auf der richtigen Ebene platziert ist (denkbar wäre auch auf Ebene der Datenfelder oder Information Concepts).
Lösungsvorschläge	Siehe Lösung

Darstellung der Lösung	Da derzeit kein Anwendungsfall für die Nutzung von Vertraulichkeitsstufen existiert, wird es aktuell nicht aufgenommen. Sofern der Bedarf dafür entsteht, kann es jedoch hinzugefügt werden. Gegebenenfalls muss diese Anforderung durch eine Änderungsanforderung an CCCEV auch für weitere Klassen hinzugefügt werden.
-------------------------------	--

Titel	Metadaten der Entitäten
Beschreibung	Es soll möglich sein im Graphen bestimmte Metadaten wie z.B. „Datum der letzten Änderung des Eintrages“ oder „Autor des Eintrages“ zu ergänzen. Diese dienen der besseren Pflege und Bearbeitung des Wissensgraphen. Damit soll es möglich sein, dass im Wissensgraph auch eine historische Entwicklung (z.B. hinsichtlich der für ein Verwaltungsverfahren notwendigen Nachweisdaten) erkennbar ist.
Lösungsvorschläge	Siehe Lösung
Darstellung der Lösung	Zum derzeitigen Stand werden keine Metadaten geführt. In den verwendeten EU-Vokabularen werden keine adäquaten Metadaten vorgegeben. Eine grobe Klärung über die notwendigen Meta-Informationen unter Einbeziehung existierender Metadaten-Beschreibungsschemata ist erfolgt. Aufgrund der Natur des Modells ist es jedoch zu jedem Zeitpunkt offen, diese Informationen hinzuzufügen.

Titel	DSMeld 19 Release
Beschreibung	Zum 01.11.2024 ist die 19. Version von DSMeld gültig. Im Projekt wurde die bei Projektstart allein vorliegende Version 18 als Grundlage verwendet.
Lösungsvorschläge	Prüfen, wo die Unterschiede in den beiden Versionen liegen.
Darstellung der Lösung	Bisher ungelöst. Dies muss in Aktualisierungen und Nachfolgeprojekten behandelt werden.

Titel	DSMeld-Blatt: Art der Wohnung
Beschreibung	Das DSMeld-Blatt 1213a „Art der Wohnung“ kann z.B. Einträge wie „Wohnung in Deutschland, aus der der Einwohner eingezogen ist“ oder „Wohnung des Ehegatten/Lebenspartners“. Diese Angabe erschließt sich sinnhaft nur in Verbindung mit den Adressendaten (Straße, Hausnummer, PLZ, Ort)
Lösungsvorschläge	Zunächst keine weitere Beziehung erstellen, sondern im „Information Requirement“ auf alle Datenfelder einmal verweisen.

Darstellung der Lösung	Bisher ungelöst. Dies muss in Aktualisierungen und Nachfolgeprojekten gelöst werden. Eventuell ist die Frage unter dem in „5.2 Abhängigkeitslogik der DSMeld Blätter untereinander“ genannten Argument ohne Relevanz.
-------------------------------	---

Titel	Umgang mit Ähnlichkeitsbeziehungen
Beschreibung	Datenfelder wie z.B. „Vorname“ kommen in verschiedenen (Verwaltungs-) Bereichen vor, haben jedoch die gleiche Bezeichnung. Es gilt zu eruieren, wie damit im Modell umgegangen wird.
Lösungsvorschläge	Eine Möglichkeit ist die Verwendung von Konzepten, um Ähnlichkeitsbeziehungen zwischen verschiedenen Datenfeldern anzugeben, wie „etwas ist gleich“, „etwas ist ähnlich“, „etwas ist ähnlich und hat weiter/kürzer reichende Aussagekraft“, „etwas ist ungleich“ o.ä.
Darstellung der Lösung	<p>Die benötigten Konzepte sind aktuell in verschiedenen Frameworks (skos; OWL) vorhanden. An dieser Frage hängt eine weitreichendere Anpassung des Modells, zudem ist auf Basis der aktuellen Instanzdaten noch unklar, welche Arten von Aussagen möglich sein müssen.</p> <p>Diese Thematik muss jedoch bei nachfolgenden Anpassungen und Erweiterungen des Modells berücksichtigt werden.</p>

Titel	Semantisch nachnutzbare Modellierung von Rechtsgrundlagen
Beschreibung	<p>Im Rahmen von Verwaltungsverfahren spielen Rechtsgrundlagen eine wichtige Rolle. Aktuell werden diese innerhalb des Projektes mithilfe eines Textverweises auf gesetz-im-internet.de dargestellt, einer Webseite des Bundesjustizministeriums, welche Gesetze veröffentlicht.</p> <p>Zugleich existieren jedoch inzwischen Vokabulare, um Gesetze semantisch nachnutzbar darzustellen wie z.B. der European Legislation Identifier (ELI). Diese erlauben auch eine genaue semantische Darstellung der Gesetze bis hin zum entsprechenden Paragraphen und Absatz. Für den deutschen nationalen Kontext liegen allerdings noch keine strukturiert in ELI abgebildeten Gesetze vor, auf die verwiesen werden könnte für die modellierten Fälle.</p>
Lösungsvorschläge	Die Elemente und Normierung deutscher Gesetze von ELI sollten im Wissensgraphen inkludiert und berücksichtigt werden.
Darstellung der Lösung	Eine Implementierung auf Basis von ELI ist zum jetzigen Zeitpunkt nicht sinnvoll. Eine zukünftige Anpassung und Erweiterung des Modells sollten dies jedoch berücksichtigen. Aktuell sind die Rechtsgrundlagen am Information Requirement entsprechend der CCCEV-Vorgaben als String modelliert. Denkbar wäre zukünftig auch eine Verknüpfung der Rechtsgrundlagen mit

	dem Public Service. Hier sieht CPSV eine Angabe auf Basis von ELI bereits vor.
--	--

4.5 Abweichungen von und Änderungsvorschläge zu den EU Core Vocabularies

4.5.1 Abweichungen vom Kernmodell CCCEV

Das Grundmodell von CCCEV enthält die grundlegende und für RegCheck nutzbare Logik einer Nachweisanforderung, die durch einen Nachweis bedient wird. Wie sich nicht zuletzt durch Studium der Webinare³⁰ aus der Entwicklungsgeschichte des Datenmodells ergeben hat und durch wenige Anwendungsbeispiele illustriert war, funktionierte diese Logik zunächst einmal auf der Ebene von Nachweisen. In diesem Kontext war der gesamte „Strang“ des Modells mit EvidenceTypes und EvidenceTypeLists (über die z.B. hätte festgelegt werden können, dass entweder ein Personalausweis oder ein Reisepass + Meldebestätigung denselben Sachverhalt belegen) nicht notwendig. Die gesamte für RegCheck im aktuellen Stadium nötige Logik konnte über Information Requirements als Anforderungsbündel, die auf InformationConcepts als einzelne Datenfelder zeigen, abgebildet werden.

Weiterhin enthält das Grundmodell eine Logik, mit der nicht nur spezifische Daten abgefragt, sondern auch Fragen im Sinne des Registermodernisierung-Reifegrads D2 gestellt werden können (statt eines Geburtsdatums z.B. die Frage „Ist die Person volljährig?“). Hierzu sind die Subklassen „Criterion“ und „Constraint“ sowie „SupportedValue“ von Bedeutung. Aufgrund der Tatsache, dass in der Registermodernisierung und für dieses Projekt prioritär der Reifegrad D1 von Relevanz ist, sind die entsprechenden Klassen vorerst außen vor geblieben.

³⁰ Webinare von DIGIT SEMIC community: <https://interoperable-europe.ec.europa.eu/collection/semic-support-centre/webinars-and-awareness-raising-sessions>

Grundmodell von CCCEV 2.1.0 (Streichungen in Rot)

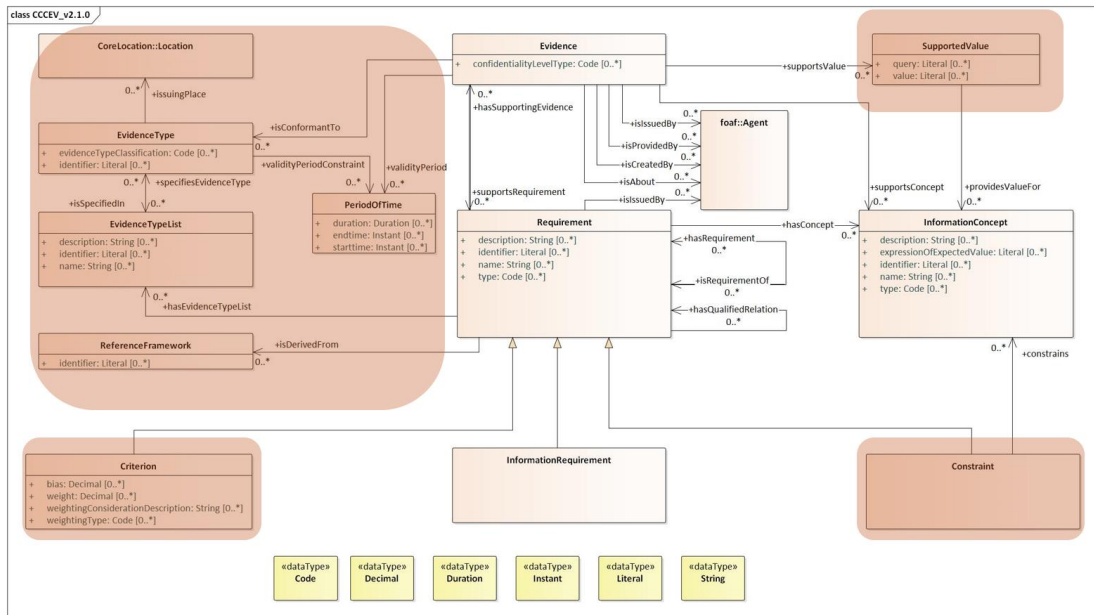


Abbildung 16: Grundmodell von CCCEV 2.1.0

Übersicht über das aktuelle RegCheck-Datenmodell

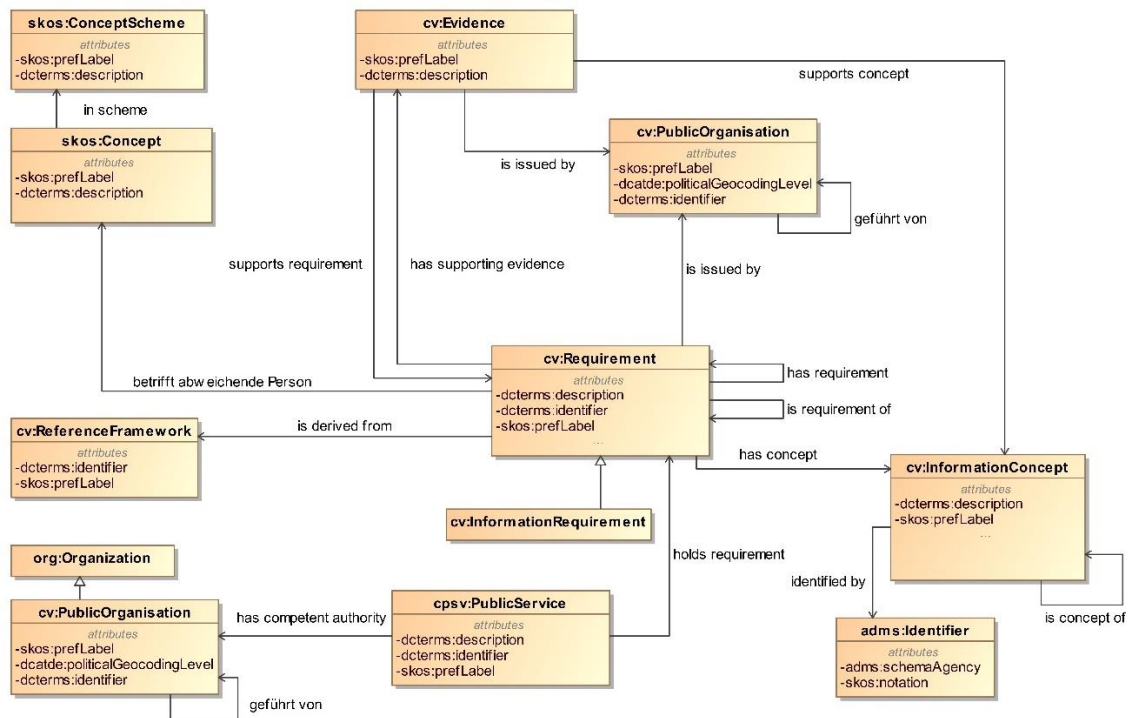


Abbildung 17: Das CCCEV-AP.de (Regcheck) Datenmodell

4.5.2 Änderungsvorschläge an CCCEV

Parallel zu den projektbezogenen Anpassungen am Modell wurde eine Reihe von potenziell wertvollen Änderungen am Grundmodell erarbeitet, die in Form von Issues an die Editoren der EU Core Vocabularies im SEMIC Support Centre der EU-Kommission weitergeleitet wurden. Diese sollen als eine mögliche Grundlage dienen für eine Erweiterung bzw. Anpassung des CCCEV Vokabulars basierend auf den praktischen Erfahrungen aus dem RegCheck Projekt, das die Logik des Datenmodells erstmalig auf Datenfeld- statt auf Nachweisebene angewandt hat.

Ticket 65 – Reifegrad-Taxonomie für Nachweise benötigt³¹

IT-PLR Beschluss <https://www.it-planungsrat.de/beschluss/beschluss-2024-15> wird ein Reifegradmodell für Nachweise beschlossen:

Dieses kennt die Codes

A, B, C, D1 und D2

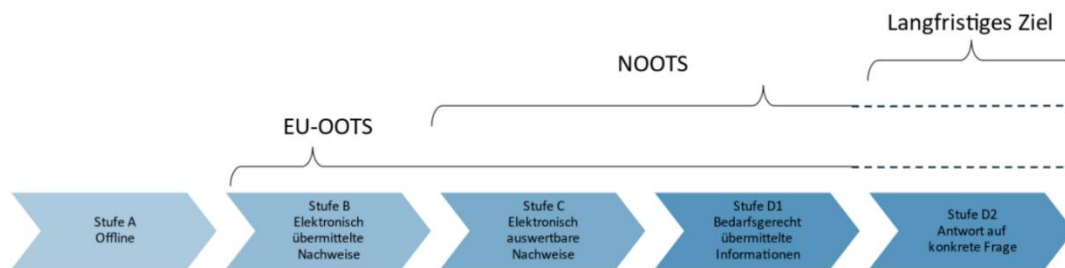


Abbildung 18: RegMo Reifegrad 2.0

Auf EU-Ebene ist bei SDG folgendes bekannt

Angaben von Mediatype (PDF, XML, SVG, PNG?) steuert die Aussage, ob ein Nachweis strukturiert ist oder nicht.

https://code.europa.eu/oots/tdd/tdd_chapters/-/blob/master/OOTS-EDM/code-lists/OOTS/OOTSMediaTypes-CodeList.gc

Der Mediatype PDF wurde eingestuft als "unstructured" (was inhaltlich auch falsch ist, da es auch in PDF befindliches XML geben kann)

Es fehlt auch die Möglichkeit zu differenzieren, dass ein Nachweis nicht nur strukturiert vorliegt (C), sondern ein besonderes Augenmerk auf die bedarfsgerechte datenminimierende Ausgestaltung des Nachweises gelegt wurde (D1). Auch können Nachweise in CCCEV-AP

³¹ <https://github.com/SEMICeu/CCCEV/issues/XY>.

nicht als "analoge" Nachweise (A) markiert werden. (Was den Übergang in die Digitalisierung aber vereinfachen könnte).

Es gibt auf EU-Ebene keine Möglichkeit, mit dem CCCEV-AP Vokabular den Reifegrad so strukturiert auszudrücken, wie das gerade in Deutschland für die Registermodernisierung möglich ist.

Daher soll das Thema auf EU-Ebene eingebracht werden:

- a) Anregen eine eigene Property dafür vorzusehen in der EvidenceType Klasse und
- b) einen Authority Table maturity level beim Publications Office (mit Vocbench angelegt?) mit den oben skizzierten Werten zu bekommen und diese zunächst auf Deutsch und Englisch zu hinterlegen.

Ticket 59 - Möglichkeit zur Modellierung von Anträgen benötigt³²

Eine Analyse im Projekt ergab, dass ein Bedarf für die korrekte Darstellung und Modellierung von Anträgen vorhanden ist. CCCEV-AP dient zur Darstellung von Nachweisen im Verwaltungskontext und CPSV zur Darstellung von Onlinediensten. Anträge von Bürgern oder Unternehmen sind relevante Verwaltungsprozessbestandteile, welche sich aktuell mit den Klassen von CCCEV-AP und CPSV nicht darstellen lassen. Daher wäre eine Erweiterung von CCCEV um diese Bestandteile oder ein neues Core Vocabulary sinnvoll. Im Rahmen des Projektes wurde dies nicht umgesetzt, jedoch wurde für eine zukünftige Erweiterung des Projektes ein entsprechendes Issue eröffnet.

Ticket 66 - Musternachweise verlinken können in Evidences³³

Es existieren oft Links auf beispielhafte Nachweise. So ist z.B. eine Geburtsurkunde beispielhaft in der Personenstandverordnung abgebildet. Auch bei SDG wird die Arbeit der Unterarbeitsgruppe „Evidence Mapping“ durch gemeinsamen Blick auf pseudonymisierte Nachweise unterstützt. Der Bedarf mit einer Property diese „exampleURL“ ausdrücken zu können, wurde als Issue 66 gemeldet.

Ticket 62 - Elemente zur Modellierung von Datenabrufen für Nicht-Hauptakteure benötigt³⁴

Es wurde im Projekt erkannt, dass im Rahmen eines Verwaltungsverfahrens häufig Informationen von Akteuren abgefragt werden, die nicht Hauptakteure des Prozesses sind, wie z.B. Partner oder Kinder des eigentlichen Akteurs, zu dem Informationen abgefragt werden. Aktuell lässt sich dies nicht ideal im Rahmen von CCCEV-AP oder CPSV modellieren. Es wurde daher,

³² <https://github.com/SEMICEu/CCCEV/issues/59>.

³³ <https://github.com/SEMICEu/CCCEV/issues/66>.

³⁴ <https://github.com/SEMICEu/CCCEV/issues/62>.

wie beschrieben, eine eigene Lösung in Form von Konzeptlisten gefunden. Nichtsdestotrotz wäre aus Gründen der besseren Nachnutzung eine modellinhärente Lösung in CCCEV empfehlenswert. Aufgrund dessen wurde ein Issue eröffnet.

Ticket 61 - Elemente zur Modellierung von Nutzungszwecken benötigt³⁵

In den Gesprächen mit den Mitarbeitenden in der Verwaltung zeigte sich, dass der Nutzungszweck eines Datenfeldes innerhalb des Verwaltungsprozesses eine relevante Information für die Mitarbeitenden ist. Aktuell kann weder CCCEV noch CPSV dies darstellen. Wie bereits beschrieben, wurde eine eigene Lösung im Projekt entwickelt. Da jedoch basierend auf den Gesprächen mit Verwaltungsmitarbeitenden davon auszugehen ist, dass dieses Thema ein allgemein in der Verwaltung wichtiger Aspekt darstellt, wurde diese Information in Form eines Issue an die Editoren des Vokabulars weitergegeben.

Ticket 63 - Zusätzliche Metadaten benötigt³⁶

Für die Bearbeitung und Pflege des Wissensgraphen sind Metadaten von Vorteil, welche verschiedene zusätzliche Informationen zu den hinterlegten Inhalten darstellen. Dazu gehören zum Beispiel Verfasser eines Eintrags, zeitliche Gültigkeit eines Eintrags etc. Einige der Attribute werden in BRegDCAT-AP durch ADMS bereitgestellt. Es stellt sich die Frage, ob auch CCCEV von den Möglichkeiten des ADMS-Vokabulars profitieren können sollte. Auch dieses Erkenntnis wurde in Form eines Issue gemeldet.

Ticket 56 - „description“ Eigenschaft benötigt³⁷

Eine Eigenschaft „description“ würde eine wertvolle Ergänzung im CCCEV Modell darstellen. Damit wäre es möglich, menschenlesbare Informationen im Modell zu hinterlegen und sowohl die praktische Nutzung als auch Bearbeitung des Modells zu erleichtern. Im Rahmen des Projektes wurde hierfür auf „dcterms:description“ aus dem Dublin Core Vocabulary zurückgegriffen.

Ticket 57 - „skos:prefLabel“ Feld benötigt³⁸

Im Projekt wurde herausgearbeitet, dass ein Feld „skos:preflabel“ für eine Reihe von Klassen notwendig ist. Die Bezeichnungen der Klassen innerhalb des Projektes folgen einer internen eindeutigen Systematik, weshalb parallel die Angabe einer natürlichsprachlichen Variante notwendig ist. Die Bezeichnung der Klassen „ReferenceFramework“ und „Evidence“ werden

³⁵ <https://github.com/SEMICeu/CCCEV/issues/61>.

³⁶ ³ <https://github.com/SEMICeu/CCCEV/issues/63>.

³⁷ <https://github.com/SEMICeu/CCCEV/issues/56>.

³⁸ <https://github.com/SEMICeu/CCCEV/issues/57>.

bei der datenfeldspezifischen Nutzung in RegCheck zudem rechtlich gefasst und abgekürzt. Ein „skos:preflabel“ erlaubt eine natürlichsprachliche, nicht abgekürzte Variante, die für die Nutzung unabdingbar ist und die Bearbeitung erleichtert. Im Rahmen des Projektes wurde hierfür auf dieses Element zurückgegriffen, was jedoch nicht originär Teil von CCCEV ist, sondern aus dem SKOS (Simple Knowledge Organization System) -Vokabular stammt. Ein CCCEV-inhärentes Element erscheint somit sinnvoll.

5 Kostenkalkulation

Auf Basis der im Projekt gewonnenen Erkenntnisse und Ausbauperspektiven lässt sich der erstellte Wissensgraph grundsätzlich in einer größeren Dimension aufsetzen. Wir empfehlen, als potenzielles Folge- und Anschlussprojekt einen agilen Projektansatz, um die Potentiale von RegCheck weiter zu vertiefen. Die Folgeprojektausrichtung kann dabei die in den Ausbauperspektiven des Gutachtens skizzierten Potentiale, z.B. RegCheck als Datenbackbone eines Verwaltungs-Chatbots, RegCheck als Auskunftssystem für die Verwaltung oder RegCheck als Adapter für EU-OOTS Daten und Datenräume in der EU, aufnehmen.

Grundsätzlich kann bei einem vollumfänglichen Folgeprojekt von einer Laufzeit über zwei Jahre ausgegangen werden. Dabei folgt der 6-monatigen Konzeptions- und Datenkonvertierungsphase die Umsetzung der Lösung sowie die Schaffung weiterer Schnittstellen.

Im Folgenden wird ein kurzer Überblick darüber gegeben, welcher personellen Ressourcen es bedarf, um eine beispielhafte Umsetzung mit den skizzierten Zielgruppen und Anwendungsszenarien zu realisieren. Für die Entwicklung des Datenmodells sowie des Wissensgraphen werden die nachfolgenden aufgelisteten Profile benötigt. Neben der fachlichen Erarbeitung wird dabei die technische Umsetzung mit entsprechendem Personal hinterlegt.

In dem „Full-Service Projekt“ werden folgende Gewerke gesehen:

- Projektleitung
- Semantic Web Consulting
- User Experience Design (UX)
- Data Analyst
- Fachkonzeption
- (Java) Entwicklung

Bei einer Laufzeit über zwei Jahren und einer durchschnittlichen Teambesetzung von ca. 14 Mitarbeitern wird ein Finanzvolumen über ca. 2.700 Personentagen benötigt. Diese Annahmen beruhen auf Erfahrungswerten, d.h. sie können abweichen. Neben diesen Personentagen wird davon ausgegangen, dass zusätzlich ca. 15% Projektbetreuung aus dem Verwaltungsbereich benötigt werden. Dies würden ungefähr 400 PT entsprechen. Eine Unterstützung durch Mitarbeitende der Verwaltung würde auf mehreren Ebenen benötigt. Neben Personalkosten werden hier Sachkosten kalkuliert, um Reise zu Vernetzungs- und Austauschtreffen zu ermöglichen. Diese müssen ebenfalls für den Zeitraum von zwei Jahren betrachtet werden.

In der Annahme, dass für die Lösungserstellung ein Hosting in einem BSI zertifizierten Rechenzentrum mit deutschsprachigem Support benötigt wird, kann davon ausgegangen werden, dass die Betriebskosten für eine Server Infrastruktur bei jährlich ca. 100.000 € liegen. Die Zahl kann abweichen und beruht auf ersten Erfahrungs- und Schätzwerten. Bei einem

internen Betrieb innerhalb eines Landesrechenzentrums fallen keine externen Betriebskosten an. Üblicherweise fallen bei der Nutzung von Open Source Software keine Lizenzkosten an. Auf Ebene der Hardware müssen hier jedoch Kostenpunkte für performante Hardware zum sicheren Betrieb des Wissensgraphen vorgesehen werden.

Bitte beachten Sie, dass es sich bei den aufgeführten Schätzungen lediglich um eine sehr grobe Kostenschätzung, welche auf Basis des aktuellen Wissens, den stattgefundenen Gesprächen und den daraus abgeleiteten Erfahrungswerten fundiert, handelt. Die tatsächlichen Aufwände können im Projektverlauf nach oben oder unten abweichen.

6 Fazit

Zusammenfassend ist festzuhalten, dass mit dem Gutachten

- a) Einblick in erhaltene Erkenntnisse gegeben
- b) Offene Fragestellungen dokumentiert
- c) Geklärte Fragestellungen aufgelistet
- d) ein Ausblick auf einen Kostenansatz für einen weiteren Ausbau des Projektes hin zu einem „Wissensgraphen der deutschen Verwaltung“ gegeben

werden konnte.

Die Autoren des Gutachtens unterstreichen die Vielzahl an gewonnenen Erkenntnissen und die positive Grundtendenz, dass sich einige Herausforderungen im Bereich der semantischen Interoperabilität sehr gut – mit dem von der EU praktizierten Ansatz des Einsatzes von „Semantic Web Technology“ bewerkstelligen lassen.

Es wird an dieser Stelle darauf hingewiesen, dass die Autoren während der gesamten Projektlaufzeit im Austausch mit Fachkollegen standen. In Folge des stetigen Austausches konnten sowohl neueste Erkenntnisse einfließen und Perspektivwechsel vorgenommen werden, als auch das Projekt präsentiert werden. RegCheck wurde auf Tagungen und Konferenzen vorgestellt, eine Vernetzung erfolgte national sowie international. In der zu Projektbeginn erarbeiteten Umfeldanalyse wird die Beziehung zu fachlich verwandten Projekten vertiefend hergestellt. Diese Analyse wurde zum Ende der Projektlaufzeit aktualisiert.

Es wird für eine Annäherung des aktuell in Deutschland gefahrenen klassisch-konservativen Ansatzes (relationale Datenbanken, FDK nur für bereits harmonisierte Dinge, keine Schnittstelle zu Registerlandkarten der EU) an diesen im Projekt erfolgreich erprobten Ansatz der „Datenhaltung in einem Wissensgraph“ plädiert.

6.1 Glossar

Begriff	Beschreibung/ Erläuterung
DSMeld	Einheitlicher Datensatz für das Meldewesen: Der Datensatz für das Meldewesen (Einheitlicher Bundes-/Länderteil) – DSMeld – ist auf Grundlage des Melde-rechtsrahmengesetzes erstmals am 21. Oktober 1982 von der Bundesvereinigung der kommunalen Spitzen-verbände herausgegeben worden. Mit Inkrafttreten des Bundesmeldegesetzes am 1. November 2015 wird die Herausgeberschaft des DSMeld – gestützt auf die Regelungen des Bundesmeldegesetzes – auf die Koor-dinierungsstelle für IT-Standards (KoSIT) übertragen.
FIM	Das Föderale Informationsmanagement (FIM) dient dazu, leicht verständliche Bürgerinformationen, ein-heitliche Datenfelder für Formulare Systeme und stan-dardisierte Prozessvorgaben für den Verwaltungsvoll-zug bereitzustellen. Ziel ist es, den Übersetzungs- und Implementierungsaufwand rechtlicher Vorgaben zu senken. Länder und Kommunen sollen - bezogen auf die redaktionelle und organisatorische Umsetzung ei-nes Verfahrens - nicht mehr für sich allein agieren müssen. Stattdessen können sie auf qualitäts-gesicherte Vorarbeiten der nächsthöheren Verwal-tungsebene zurückgreifen.
Nachweise	Nachweise sind alle Unterlagen oder Daten, ein-schließlich Text- oder Ton-, Bild- oder audiovisuellen Aufzeichnungen, unabhängig vom verwendeten Me-dium, die von einer zuständigen Behörde verlangt werden, um Sachverhalte nachzuweisen oder die Ein-haltung von Verfahrensvorschriften zu belegen. Nach-weise können sowohl als analoge („Unterlagen“) wie auch digitale Informationen („Daten“) vorliegen. Ein Nachweis kann z.B. eine Geburtsurkunde sein. Siehe auch: Evidence.
Nachweistyp	Nachweistypen dienen zur Klassifikation von Nach-weisen nach gemeinsamem Zweck oder Inhalt. Nach-weistypen sind selbst keine Nachweise, aber Nach-weise gehören zu einem Nachweistyp. Beispiel: Alle konkreten Geburtsurkunden sind Aus-prägungen zum Nachweistyp "Geburtsurkunde".
NOOTS	Das National-Once-Only-Technical-System (NOOTS) ist ein System aus technischen Komponenten, Schnittstellen und Standards sowie organisatorischen und rechtlichen Regelungen, das öffentlichen Stellen

	den rechtskonformen Abruf von elektronischen Nachweisen aus den Registern der deutschen Verwaltung ermöglicht. Es basiert auf dem Gedanken des Once-Only-Prinzips.
Once-Only-Prinzip	Ein wesentliches Ziel ist bei der Modernisierung der Verwaltungslandschaft, dass die Bürger in Zukunft ihre Daten und Nachweise nicht immer wieder erneut vorlegen müssen, um Verwaltungsleistungen zu erhalten, sondern - wenn sie dem eingewilligt haben - nur einmal.
Register	Ein Register ist ein Datenbestand der öffentlichen Verwaltung. In den Registern liegen die Nachweise, welche für Verwaltungsverfahren notwendig sind.
Registermodernisierung	<p>Das Programm Registermodernisierung (RegMo) ist eines der größten Projekte im Rahmen der Digitalisierungsbestrebungen von Bund, Ländern und Kommunen.</p> <p>Einheitlich gestaltete, inhaltlich aktuelle, vernetzte Register stellen einen wichtigen Meilenstein dar für eine digitale, bürokratiearme und serviceorientierte Verwaltung, die Bürgerinnen und Bürger sowie Unternehmen entlastet.</p> <p>Ein wesentliches Ziel ist, dass Bürgerinnen und Bürger in Zukunft ihre Daten und Nachweise nicht immer wieder erneut vorlegen müssen, um Verwaltungsleistungen zu erhalten, sondern - wenn sie dem eingewilligt haben - nur einmal (Once-Only-Prinzip).</p>
Registertyp	<p>Registertypen dienen zur Klassifikation von Registern nach gemeinsamem Zweck oder Inhalt. Registertypen sind selbst keine Register, aber Register gehören zu einem Registertyp.</p> <p>Beispiel: Alle Registerinstanzen der Melderegister sind Ausprägungen zum Registertyp "Melderegister".</p>

6.2 Abkürzungsverzeichnis

Abkürzung	Ausgeschrieben
FIM	Föderales Informationsmanagement
NOOTS	National-Once-Only-Technical-System
RDF	Ressource Description Framework
SDG-VO	Single Digital Gateway Verordnung

XÖV	XML für die Öffentliche Verwaltung
-----	------------------------------------

7 Anhang

7.1 Geklärte Fragestellungen

1. Welche Rollen vorzusehen sind beim Wissensgraphen? KG Builder, KG Analyst, KG User.
2. Warum man nicht an einem Tool kleben sollte? (Wechselbedarf, Proprietäre Angaben vermeiden)
3. Warum es verschiedene Ebenen des Modells gibt? (eigentliche Semantische Ebene, Modellierungsebene RDF, Validierungsebene SHACL)
4. Wieso die Softwarelandschaft eher voll funktionale Silos hat statt leicht verknüpfbare Open Source Bibliotheken?
5. Was der Unterschied ist zwischen BregDcat-AP und CCCEV?
6. Warum ein zentraler Sparql-Endpoint ausserhalb von Tools positioniert sein sollte?
7. Warum es keine ""semantische Fachdatenavigation"" geben kann? (Ende-zu-Ende lässt keinen Einblick in die Fachdaten; Hosting im BVA / ITZ Bund erlaubt keine experimentelle Software)
8. Warum es hier kein NOOTS-Integrationskonzept geben konnte?
9. Warum ein Wissensgraph so gut Fachlichkeit repräsentieren kann?
10. Wieso ein Wissensgraph mehr ist als ein Datengraph?
11. Weshalb skos und OWL deutschen Föderalismus gut abbilden kann? (Mappings, Grautöne, Mehrsprachigkeit, Hierarchien, Open World Paradigma)
12. Weshalb man manche Angaben nur in Interviews erfährt? (Wissen ist implizit vorhanden und muss "rausgekitzelt" werden)
13. Wieso man Interviews vor Ort durchführen sollte? (Interviewte brauchen den normalen Arbeitskontext)
14. Wie man das Thema Wissensgraph kommunizieren kann?
15. Welcher Use Case in der Regmo vom Link Prediction profitiert?
16. In welches Umfeld ein Nachfolgeprojekt stößt? (Regierungsbildung, unklarer Haushalt, NOOTS-Staatsvertrag)
17. Wie ein Nachfolgeprojekt aufgesetzt sein sollte?
18. Wieso ein Wissensgraph Haluzinationen einer KI abmildert?
19. Wie das Verhältnis von RegCheck zum FDK Vo.4 ist?
20. Wie langfristig Änderungen an EU-Vokabularen eingereicht werden müssen?

21. Welche fachlichen Use Cases die Meldebescheinigung brauchen?
22. Wo man KI in der RegMo einsetzen könnte?
23. Wie man Changes in die EU einbringt?
24. Wann ein Applikationsprofil zum Wissensgraph wird?