

HENRY

Hydraulic Engineering Repository

Ein Service der Bundesanstalt für Wasserbau

Conference Paper, Published Version

Huang, Jiuru

Datenabfrage und -integration im Kontext vom Multiprojektmanagement im Verkehrswasserbau mit Semantic Web Technologie

Verfügbar unter/Available at: <https://hdl.handle.net/20.500.11970/108246>

Vorgeschlagene Zitierweise/Suggested citation:

Huang, Jiuru (2021): Datenabfrage und -integration im Kontext vom Multiprojektmanagement im Verkehrswasserbau mit Semantic Web Technologie. In: Disser, Michael ; Hoffmann, André ; Kuhn, Luisa ; Scheich, Patrick (Hg.): 32. Forum Bauinformatik 2021. Darmstadt: Universitäts- und Landesbibliothek Darmstadt. S. 157-165.

Standardnutzungsbedingungen/Terms of Use:

Die Dokumente in HENRY stehen unter der Creative Commons Lizenz CC BY 4.0, sofern keine abweichenden Nutzungsbedingungen getroffen wurden. Damit ist sowohl die kommerzielle Nutzung als auch das Teilen, die Weiterbearbeitung und Speicherung erlaubt. Das Verwenden und das Bearbeiten stehen unter der Bedingung der Namensnennung. Im Einzelfall kann eine restriktivere Lizenz gelten; dann gelten abweichend von den obigen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Documents in HENRY are made available under the Creative Commons License CC BY 4.0, if no other license is applicable. Under CC BY 4.0 commercial use and sharing, remixing, transforming, and building upon the material of the work is permitted. In some cases a different, more restrictive license may apply; if applicable the terms of the restrictive license will be binding.



Datenabfrage und -integration im Kontext vom Multiprojektmanagement im Verkehrswasserbau mit Semantic Web Technologie

Jiuru Huang

Bundesanstalt für Wasserbau, Abteilung Bautechnik, Referat Infrastrukturmanagement

Zusammenfassung

Im Bereich des Verkehrswasserbaus stehen aktuell viele teilweise insulare IT-Systeme mit heterogenen Daten zur Verfügung. Für die nutzerfreundliche Informationsbereitstellung ist die Vernetzung der Daten von hoher Bedeutung. Mit Hilfe der Semantic Web Technologie können die Datenintegration der vorhandenen heterogenen Datenquellen und das Auffinden von Daten verbessert werden. In einem aktuellen Projekt „Multiprojektmanagement der Wasserstraßen- und Schifffahrtsverwaltung“ wird ein Informationssystem entwickelt, das für einen bestimmten Informationsbedarf relevante Daten aus vorhandenen IT-Systemen des Verkehrswasserbaus exportiert, mit zusätzlichen Informationen ergänzt und miteinander verknüpft. Das Ziel ist den Nutzern einen entsprechenden Überblick über die relevanten Informationen im Kontext des Multiprojektmanagements von verschiedenen Baumaßnahmen zu verleihen. Die Bündelung der Daten erfolgt aktuell noch überwiegend manuell und nicht automatisch. Mittels der Semantic Web Technologie wird für die Anwendungsfälle aus diesem Kontext ein Datenmodell mit Domainwissen aufgebaut. Dafür werden spezifische und semantische Abfragen konzipiert und die Ergebnisse visualisiert. In diesem Beitrag wird exemplarisch das vorläufige Ergebnis des Datenmodells und der semantischen Abfragen vorgestellt, um somit das Potenzial der Semantic Web Technologie in der Datenintegration im Bereich des Verkehrswasserbaus aufzuzeigen.

Keywords: Semantic Web Technologie, Verkehrswasserbau, Datenintegration, Domainwissen, Semantische Abfrage

1 Einleitung

Die Digitalisierung bietet für Planen, Bauen und Betreiben der Wasserstraßeninfrastruktur großes Potenzial. Dies zeigen auch die Empfehlungen zur Erstellung eines Masterplans Digitalisierung im Bereich „Planen, Bauen und Betreiben der Wasserstraßeninfrastruktur“ auf [1]. Die fünf wichtigsten Maßnahmen zielen darauf ab, wie die nutzerorientierte Informationsbereitstellung optimal gewährleistet werden kann. Dazu gehören z. B. die Standardisierung der Prozesse und eine einheitliche Datengrundlage [2], deren Notwendigkeit und der daraus resultierende Entwicklungsbedarf durch das vergangene, abgeschlossene Teilprojekt des Pilotprojekts der Wasserstraße- und Schifffahrtsverwaltung des Bundes (WSV) zur Erprobung von Methode Building Information Modeling (BIM) belegt wurden. Die wichtigsten Erkenntnisse aus der Bearbeitung von diesem Teilprojekt sind:

- Eine Verbesserung der nutzerorientierten und den Bedürfnissen der WSV entsprechenden Informationsbereitstellung mithilfe des prozessorientierten Ansatzes mit der Methode „Information Delivery Manual“ (DIN EN ISO 29481 [3, 4]) kann erreicht werden.
 - Es kann bei der Informationsbereitstellung weitestgehend auf Informationen bestehender Systeme der WSV zurückgegriffen werden.
 - Allerdings wird eine Umstrukturierung der bestehenden Datenverwaltung entsprechend der Ansätze der digitalen Objektbibliothek bzw. der Semantisierung empfohlen. Dies ermöglicht die Verbesserung der Auffindbarkeit benötigter Daten in den bestehenden IT-Systemen.
-



Den Erkenntnissen zur Folge wurde ein Forschungsvorhaben in der Bundesanstalt für Wasserbau (BAW) zur Erprobung der Methode Semantic Web zur Überführung relevanter Informationen in ein Datenmodell aufgestellt. In diesem Beitrag werden Zwischenergebnisse dieses Forschungsprojektes beispielhaft erläutert.

Nach einer Darstellung der Probleme und der Motivation folgt eine kurze Analyse des Stands der Technik sowie der Erkenntnis im Bereich der Anwendung Semantic Web Technologie (SWT) im Bereich des Verkehrswasserbaus. Danach folgt die Erläuterung der Methoden-Vorgehensweise zur Anwendung der SWT, die die Untersuchungsschritte beinhaltet. In dem darauffolgenden Kapitel werden Untersuchungsergebnisse, unterteilt in Kapiteln (der Domänenbereich, die Anforderungen an das Datenmodell, das Datenmodell sowie die Datenintegration), einzeln beschrieben. Mit einem Fazit und Ausblick endet dieser Beitrag.

2 Problemstellung und Motivation

Die IT-Anwendungslandschaft in der WSV ist aufgrund ihrer langen Entwicklungshistorie technologisch heterogen. Aktuell stehen in der WSV viele Informationen in Daten unterschiedlicher Art und Qualität, innerhalb und außerhalb von unterschiedlichen IT-Systemen zur Verfügung. Dadurch wird eine Kombination und Weiterverarbeitung der enthaltenen Informationen erschwert. Die Überführung von relevanten Informationen aus diesen vorgenannten Informationsquellen in ein Datenmodell ist hierbei ein zentraler Punkt. Ein Informationsmodell, das eine effizient vernetzte IT-Landschaft zur optimalen Bedienung der Informationsanforderungen der Nutzer bündelt und integriert, existiert nicht.

Eine Vernetzung der bestehenden IT-Systeme der WSV ohne neue, redundante und insulare Datenhaltungen kann geschaffen werden. Dadurch können Informationsbedürfnisse der Prozessbeteiligten optimal (aktuell, redundanzfrei, widerspruchsfrei) befriedigt werden. Die Informationskonsumenten verlieren weniger Zeit bei der Suche nach entscheidungsrelevanten Informationen. Somit kann eine effizientere Projektbearbeitung erzielt werden. Die Informationsbereitstellung kann dadurch verbessert werden. Ferner kann diese Arbeit einen Beitrag zu dem Wissensmanagement der WSV in der methodischen Frage leisten.

3 Stand der Technik und Erkenntnis

Die SWT wird als Stand der Technik unter Verwendung von offenen, einheitlichen Standards zur Integration der heterogenen Daten angewendet. Ein bekanntes Beispiel dafür ist der Ansatz Linked Open Data, welcher das Web zu einem globalen, verteilten, semantikbasierten Informationssystem bringen soll [5]. Auch im Bereich des Planen, des Bauens und des Betriebens von cyber-physischen Produktionssystemen wird diese Methode zum Aufbau einer Wissensbasis der verschiedenen interagierenden Disziplinen erfolgreich eingesetzt [6]. Im Bereich des Verkehrswasserbaus existiert bisher jedoch keine Erfahrung zum Aufbau eines Datenmodells mit Wissensbasis durch Semantic Web.

Hitzler et al. [7] erklärt das Semantic Web wie folgt: Das Semantic Web steht für die Idee, einheitliche, offene Standards für die Beschreibung von Informationen von vornherein so zu vereinbaren, dass deren Verarbeitung durch Maschinen möglich ist. Es beschäftigt sich primär mit den Problemen, welche sich aus der Fülle der im Web präsenten Informationen ergeben, beispielsweise das Auffinden relevanter Informationen, die Integration von Informationen aus verschiedenen Quellen und das Problem des impliziten Wissens und bietet Lösungsansätze durch die Sprachen zur Repräsentation von Wissen in der Form von Ontologien. Der Begriff Ontologie ist in dem Zusammenhang als äquivalent zum Begriff Wissensbasis zu verstehen, welche Wissen einer Anwendungsdomäne repräsentiert. Die Ziele des Semantic Web sind: Finde Wege und Methoden, Informationen so zu repräsentieren, dass Maschinen damit in einer Art und Weise umgehen können, die aus menschlicher Sicht nützlich und sinnvoll erscheint. Dabei mag das entstehende Verhalten von Systemen in den Augen mancher Nutzer durchaus als intelligent erscheinen.

Das als Ziel gesetzte Datenmodell mit Wissensbasis ist also dann nützlich und sinnvoll, wenn die Interpretationen der Daten von Maschinen selbständig hergeleitet werden können, welche für Menschen als nützlich erscheinen. Dabei spielen die Informationsbedürfnisse eine große Rolle. Zu dieser Fragestellung greift dieses Forschungsvorhaben weitestgehend auf die Ergebnisse der assoziierten Projekte zurück. Hierzu werden keine Einzelheiten erläutert.

Die Informationsbedürfnisse, die verteilt in den verschiedenen IT-Systemen vorliegen, werden nun in eine Wissensbasis umgewandelt. Die Wissensbasis enthält sowohl explizites als auch implizites Wissen. Der Nutzer fand in der Wissensbasis beispielsweise Informationen über „Segmentschütz“, obwohl er nach „Wehrverschluss“ gesucht hat. D. h., die im Datensatz vorhandene Entität „Segmentschütz“ wurde mit dem Begriff „Wehrverschluss“ in Relation gesetzt. Diese Informationen oder Datengebiete (Entitäten) liegen in der WSV-IT-Landschaft jedoch unterschiedlich repräsentiert dar. Um diese Daten miteinander zu vernetzen, müssen zuerst die Bedeutungen der Daten (s. Kapitel 5.3) und die Speicherorte (s. Kapitel 5.4) analysiert werden.



Diese Analysen können nicht vollständig durch die Adaption und die Anwendung existierender und aus anderen Domänenbereichen entstandenen Wissensbasen (wie z. B. Building Product Ontology [8], ifcOWL (IFC4_ADD2_TC1) [9] und Basic European Road OTL [10]) ersetzt werden, da das Wissen durch die unterschiedlichen, fachlichen Disziplinen unterschiedlich repräsentiert wird [11] und Domainwissen nur in dem eigenen Domänenbereich repräsentiert werden kann [12]. Im Verkehrswasserbau wird z. B. eine Handlung an einem Objekt als Maßnahme deklariert, wessen Beziehung noch nicht in der vorhandenen Wissensbasis abgebildet wurde und die nur in eigenem Domänenbereich existiert. Es besteht auch die Möglichkeit, die existierende Wissensbasis in die aufgebaute Wissensbasis zu integrieren [13]. Der erste Schritt zur Anwendung der Wissensrepräsentanz mit SMT ist dennoch der Aufbau der eigenen Wissensbasis.

4 Vorgehensweise

Die Methode des Semantic Web verwendet, anders als Datenbanksysteme, prinzipiell die Formulierung von Assoziationen, die beim Menschen durch bestimmte Begriffe hervorgerufen werden. Dadurch ist es möglich, die verschiedenen Informationen der unterschiedlichen Quellen miteinander in Verbindung zu setzen. In einem weiteren Schritt können eine Wissensbasis der zusammenhängenden Informationen bzw. die Netzwerke von Informationen über logischen Relationen gebildet werden. Diese zusammenhängenden Informationen können so von den Maschinen interpretiert werden, dass Menschen beim Auffinden der Informationen geholfen wird. Nachfolgend werden die prinzipiellen Untersuchungsschritte der Methode Semantic Web aufgelistet. In dem darauffolgenden Kapitel 5 werden die vorläufigen Untersuchungsergebnisse beispielhaft erläutert.

Domänenbereich: Ein Domänenbereich, in dem die Wissensbasis aufgebaut werden soll, die dem Domainwissen entspricht, wird festgelegt. Dabei ist zu untersuchen, ob das Domainwissen existiert und in welcher Qualität (s. Kapitel 5.1 Domänenbereich – Multiprojektmanagement).

Anwendungsszenarien und Anforderungen: Es werden Abfragen, die das Datenmodell bedienen soll, als Anwendungsszenarien definiert. Auf diese Abfragen aufbauend kann ein Assoziationsnetz als Entwurf für die Wissensbasis konzipiert werden, welches die im Datenmodell befindlichen Klassen bildet. Dieses Assoziationsnetz ist temporär und wird durch den nachfolgenden Schritt (Modellierung) vollständig ersetzt. Die Erfüllung der Abfragen und der Ergebnisse der Abfragen dient dem Datenmodell und der Datenintegration wiederum als Validierungsmöglichkeit (s. Kapitel 5.2 Anforderungen an das Datenmodell).

Modellierung: Die Bestandteile der Wissensbasis (z. B. Klassen, Rollen und Individuen) werden in formalen Sprachen der Methode Semantic Web definiert. Die Definitionen werden durch eine umfangreiche Recherche der Verwaltungsvorschriften und der Regelwerke widerspruchsfrei mit Quellenangaben belegt. Die Qualität des Datenmodells wird durch die vorher definierten Anforderungen überprüft (s. Kapitel 5.3 Datenmodell).

Datenintegration: Zusätzlich zu den Erkenntnissen aus einer internen Datenlandkarte – erläutert in Schenk und Laursen [14] – werden die Quellspeicherorte der Wissensbasis und deren Datenstrukturen in den Datenbeständen der verteilten IT-Systemen analysiert, um die Durchsuchungsmöglichkeit der vorhandenen IT-Systemen über die Wissensbasis zu skizzieren. Es werden nur die Datenstrukturen, die eine Bedeutung haben, in die Wissensbasis integriert. Dabei ist es irrelevant, welche Datenrelationen in den IT-Systemen gespeichert sind, da eine Wissensbasis diese integrieren kann, damit sie von Maschinen interpretiert werden können (s. Kapitel 5.4 Integration der Datenquellen).

Diese Schritte (außer dem ersten Schritt) werden iterativ wiederholt, womit eine kontinuierliche Verbesserung des Datenmodells erreicht werden kann. Das Datenmodell wird fortlaufend auf Konsistenz und Kohärenz geprüft.

5 Untersuchungsergebnisse

5.1 Domänenbereich – Multiprojektmanagement

In einem aktuellen Projekt „Multiprojektmanagement der WSV“ (MPM) wird ein Informationssystem MPM entwickelt, das für einen bestimmten Informationsbedarf relevante Daten aus vorhandenen IT-Systemen (z. B. WinD und WSVPruf) der WSV exportiert, mit zusätzlichen Informationen ergänzt und miteinander verknüpft. Das Ziel ist, den Nutzern einen entsprechenden Überblick über die relevanten Informationen im Kontext des Multiprojektmanagements von verschiedenen Baumaßnahmen zu verleihen. Die Bündelung der Daten erfolgt aktuell noch überwiegend manuell und nicht automatisch.

Es wurden exemplarische Anwendungsbeispiele aus diesem Kontext des Informationssystems MPM als Domänenbereich festgelegt. Die Gründe dafür sind zum einen, dass das Informationssystem MPM umfangreiche Daten unterschiedlicher Art und Qualität aus



vorhandenen IT-Systemen der WSV nutzt. Zum anderen wurden die im MPM genutzten Informationen in einem Standardisierungsprozess festgelegt, wodurch ein einheitliches Verständnis der Begrifflichkeiten und somit eindeutige Definitionen erzielt wurden. Darüber hinaus sind die Nutzerbedürfnisse, festgelegt in Datenbankenkonzepten, relativ klar formuliert. Aus diesen Gründen eignet sich das Informationssystem MPM als Domänenbereich für eine Wissensbasis. Das Domainwissen ist in einer guten Qualität.

5.2 Anforderungen an das Datenmodell

Die Abfragen, die das Datenmodell mit Wissensbasis bedienen soll, wurden mit Berücksichtigung des Nutzerbedarfs im Informationssystem MPM als mögliche Anwendungsszenarien grob formuliert. Sie werden im Laufe der Datenmodellierung weiter konkretisiert und verbessert (s. Abbildung 1, oberer Bereich). Es wurde ein Assoziationsnetz (s. Abbildung 1, unterer Bereich) passend zu den Abfragen aufgestellt, bei denen die gesuchten Schlüsselwörter (mit Schriftschnitt in Fett) im Assoziationsnetz enthalten sind.

Typ der Fragen	Einfache Frage	Konkret kann es so formuliert werden
Objekt-Abfrage	Welche Wehrverschlüsse der Wasserstraße xy, haben einen schlechten Zustand ?	Abfrage: Suche nach Wehrverschlüssen, die eine Zustandsnote >3,7 haben und die an einer bestimmten Wasserstraße (oder einem Abschnitt) liegen. Ergebnis der Abfrage: Darstellung aller betroffenen Wehrverschlüsse mit Namen der Werkanlage auf einer geografischen Karte
Maßnahmen-Abfrage	Wann wurden die jeweils letzten Instandsetzungsmaßnahmen an den Wehrverschlüssen der Wasserstraße xy durchgeführt?	Abfrage: Suche nach den jeweils letzten Instandsetzungsmaßnahmen an den Wehrverschlüssen an einer bestimmten Wasserstraße (oder einem Abschnitt) Ergebnis der Abfrage: Auflistung aller dazugehörigen Instandsetzungsmaßnahmen nach Jahren und mit Namen der Wehranlage auf der Zeitskala
Dokumenten-Abfrage	Liegen alle geprüften / genehmigten Unterlagen der Voruntersuchung von Projekten mit Wehren vor?	Abfrage: Suche nach allen geprüften / genehmigten Unterlagen der Voruntersuchung von Projekten mit Wehren Ergebnis der Abfrage: Auflistung aller dazugehörigen Unterlagen

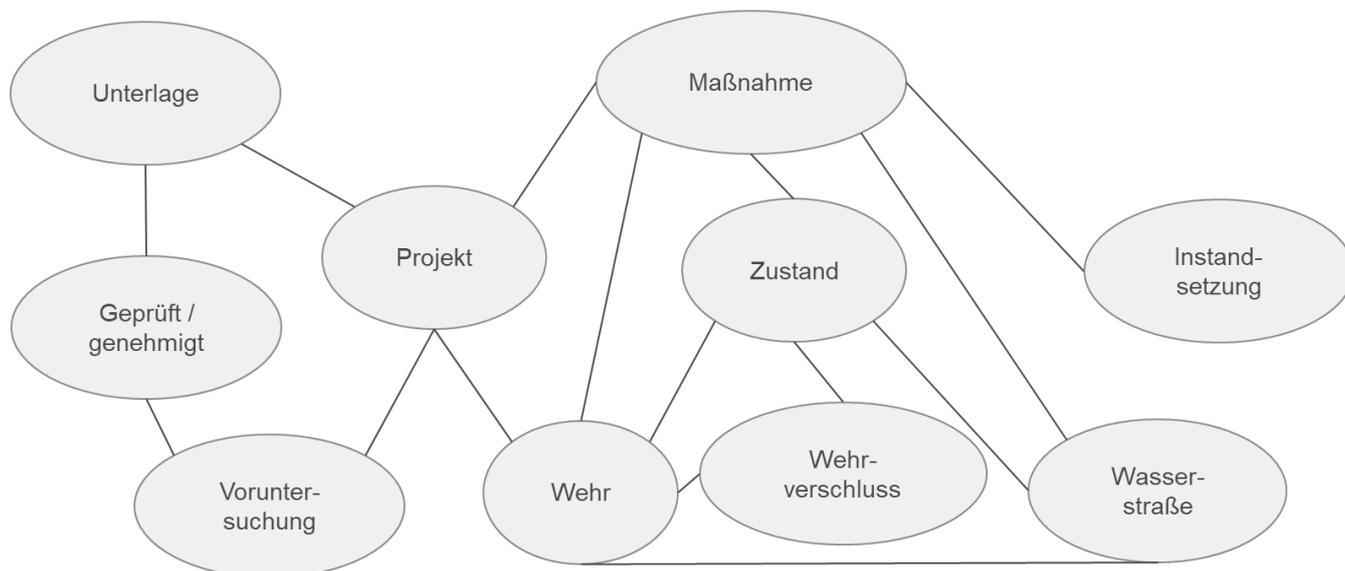


Abbildung 1: Beispiele der möglichen Abfragen und Ergebnissen der Abfragen (oben) sowie die dazugehörigen Klassen in einem beispielhaften semantischen Assoziationsnetz (unten) (eigene Darstellung)

Eine Typisierung von Abfragen mit einem typischen Frage-Beispiel kann bei der Modellierung Orientierung verschaffen, da bestimmte Entitäten in den Fragen vorkommen und diese dargestellt werden können, welche in einem Assoziationsnetz verbunden werden. Dieses Assoziationsnetz kann aufgrund von Übersichtlichkeit dem Modellierer wiederum bei der Erstellung von Klassen,

Properties und Individuen helfen. Final können dann die logischen Beziehungen formuliert werden, welche die ungerichteten Kanten ersetzen.

Im Kontext vom Informationssystem MPM wurden die folgenden zentralen Typen von Abfragen formuliert:

Objekt-Abfragen: wie z. B. „Welche Wehrverschlüsse der Wasserstraße xy haben einen schlechten Zustand?“

Maßnahmen-Abfragen: wie z. B. „Wann wurden die jeweils letzten Instandsetzungsmaßnahmen an den Wehrverschlüssen der Wasserstraße xy durchgeführt?“

Dokumenten-Abfragen: wie z. B. „Liegen alle geprüften / genehmigten Unterlagen der Voruntersuchung von Projekten mit Wehren vor?“

Ein Kriterium der Nutzerfreundlichkeit ist auch, dass die Ergebnisse dieser Abfragen für Menschen visualisiert (in Form von Listen, Tabellen, geographischen Karten, Zeitskalen, usw.) dargestellt werden können. Die Datenvisualisierung stellt jedoch kein Fokus der Arbeit dar.

5.3 Datenmodell

Der Aufbau des Datenmodells befindet sich noch in einem frühen Stadium. Es können vorläufige exemplarische Ausschnitte der Wissensbasis in Form von Graphen (Knoten: Subjekte, Kanten: Prädikate und Knoten: Objekte) präsentiert werden.

Nachfolgend werden Beispiele aus der Wissensbasis explizit erläutert.

Es wurden folgende Definitionen festgelegt:

- Wenn eine Maßnahme einen Entwurf-AU erfordert dann ist diese Maßnahme ein Projekt (Investitionsprojekt).
- Wenn eine Maßnahme im HH-Kapitel 1203 und Titel 711 01/ 712 01 /780 01/02/04/05 eingeordnet ist, dann ist diese Maßnahme eine bauliche Maßnahme (an einer Bundeswasserstraße), (ansonsten Beschaffung).

Aus den Definitionen kann ein Graph wie Abbildung 2 dargestellt werden. Eine nützliche Eigenschaft einer Wissensbasis ist es, dass zusätzliches Wissen auch implizites Wissen aus der Wissensbasis abgeleitet werden kann. Eine Erläuterung dieser Ableitung wird wie folgt beschrieben.

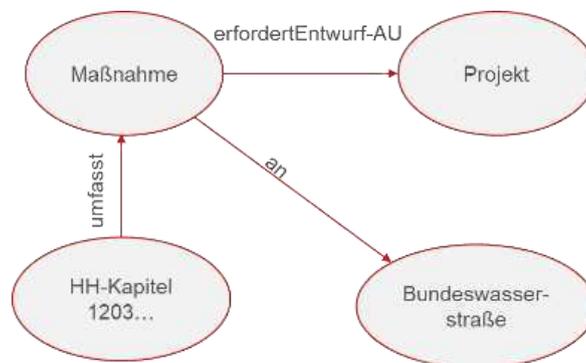


Abbildung 2: Ein Graph für Entitäten: „Maßnahme“, „Projekt“, „Bundeswasserstraße“ und „HH-Kapitel 1203...“ (eigene Darstellung)

Nach der Prädikatenlogik der ersten Stufe kann der Graph wie folgt geschrieben werden:

$$(\forall x) (\text{Projekt}(x) \rightarrow ((\exists y) (\text{erfordertEntwurf-AU}(x,y) \rightarrow \text{Maßnahme}(y))) \quad (1)$$

$$(\forall x) (\text{Maßnahme}(x) \rightarrow ((\exists y) (\text{umfasst}(x,y) \rightarrow \text{HH-Kapitel1203}(y))) \quad (2)$$

Aus dem Graph bzw. aus den Formeln 1 und 2 kann schlussgefolgt werden:

$$(\forall x) (\text{Projekt}(x) \rightarrow ((\exists y) (\text{erfordertEntwurf-AU}(x,y) \rightarrow \text{HH-Kapitel1203}(y))) \quad (3)$$

D. h.: Alle Projekte haben mind. eine Maßnahme im HH-Kapitel1203, welche einen Entwurf-AU erfordert. Umgekehrt kann nicht abgeleitet werden, dass eine Maßnahme im HH-Kapitel1203 ein Projekt ist. Dies entspricht der Inferenzregel der Transitivität.

Diese vorgestellte Wissensbasis inkludiert die Entitäten von „Maßnahme“, „Projekt“, „Bundeswasserstraße“ und „HH-Kapitel 1203...“. Ein anderes Beispiel der Wissensbasis (s. Abbildung 3), gegenüber dem vorherigen Beispiel (s. Abbildung 2), wird nachfolgend erläutert, welche einen anderen Bereich der Wissensbasis mit den „Objekten“-Entitäten einschließt.

Es wurden folgende Definitionen festgelegt:

- Wenn ein Anlagenkomplex an einer Fallstufe eines staugeregelten Flussabschnittes der Bundeswasserstraße liegt, dann ist dieser Anlagenkomplex eine Staustufe.
- Wenn ein festes Absperrbauwerk als überwiegender Teil einer Staustufe dient, dann ist dieses feste Absperrbauwerk ein Wehr.
- Wenn ein Verschluss Teil eines Wehres mit beweglichen Verschlüssen ist und er kein Revisionsverschluss ist, dann ist der Verschluss ein Wehrverschluss

Aus diesen Definitionen kann ein Graph wie Abbildung 3 dargestellt werden.

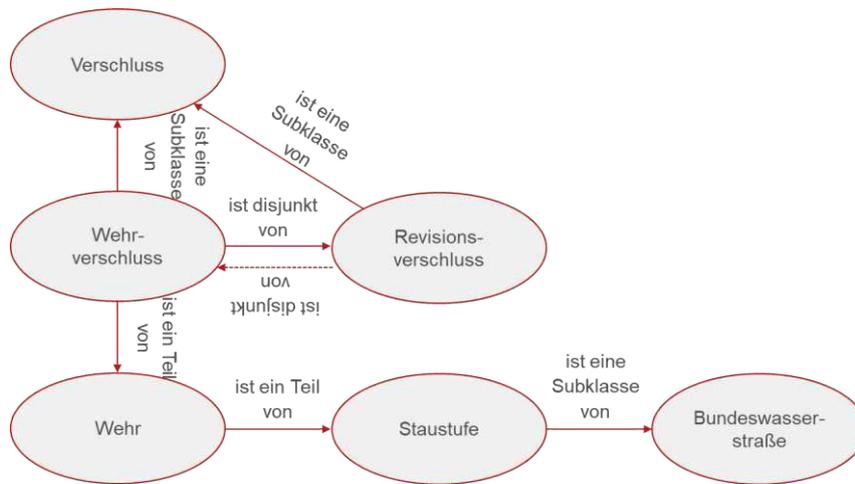


Abbildung 3: Ein Graph für die Entitäten: „Verschluss“, „Wehrverschluss“, „Revisionsverschluss“, „Wehr“, „Staustufe“ und „Bundeswasserstraße“ (eigene Darstellung)

Daraus kann – ähnlich wie in dem vorherigen Beispiel – folgendes beispielhaft abgeleitet werden:

Alle Wehre sind mind. ein Teil der Bundeswasserstraße, obwohl sich dies von den Definitionen an sich nicht explizit deklarieren lässt. Hierin entsteht ein Phänomen des impliziten Wissens, welches nur aus den Mengen der Beziehungen abgeleitet wird, wodurch es dem Nutzer bei der Suche nach implizitem Wissen helfen kann – eine nützliche Eigenschaft der Wissensbasis.

Aus diesen zwei Beispielen kann dann eine Wissensbasis nach der Web Ontology Language (OWL) gebildet werden, wobei die Grundbausteine von OWL-Ontologien aus Klassen und Properties bestehen. Die Klassen hier sind die Subjekte oder Objekte (im Graph als Knoten dargestellt) und die Properties hier sind die Prädikate (im Graph als gerichtete Kanten).

Weitere Definitionen für „Projekt“-Informationen wurden aus dem Fachkonzept Informationssystem MPM abgeleitet. Ein wesentlicher Bestandteil der Wissensbasis mit „Projekt“-Informationen ist es, dass ein „Projekt“ aus zwei wesentlichen Klassen – „Maßnahme“ und „Objekt“ – besteht, von denen unterschiedliche „Projekt“-Klassen abgeleitet werden. Dies ermöglicht die Abfrage wie z. B. „alle Projekte der Instandsetzung und/oder des Wehrverschlusses“, bei der „Instandsetzung“ eine „Maßnahme“ ist und „Wehrverschluss“ ein „Objekt“ ist.

Dabei wurden folgende Definitionen festgelegt:

- ! Maßnahmen haben Handlungen an einem Objekt mit dem Zweck ein klar definiertes Ziel zu erreichen zum Gegenstand

Aus der Definition kann ein Graph wie Abbildung 4 gebildet werden, wodurch eine Brücke von „Projekt“ zu „Objekt“ geschlagen werden kann. Dies ermöglicht eine Erweiterung der zwei vorgenannten Beispiel-Ontologien (s. Abbildung 2 und Abbildung 3).

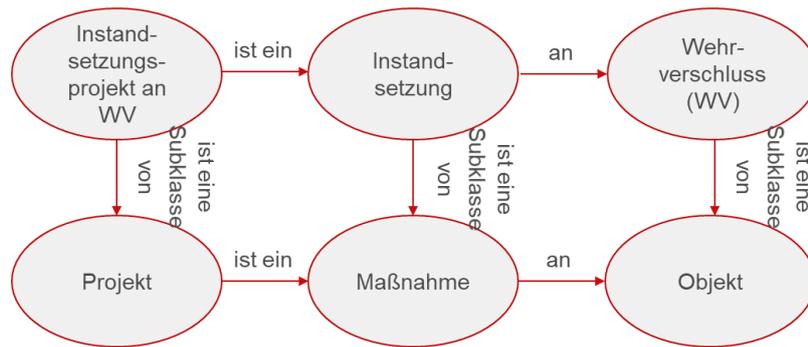


Abbildung 4: Ein Graph für die Entitäten „Projekt“, „Maßnahme“ und „Objekt“ sowie eine von ihren Subklassen (eigene Darstellung)

Des Weiteren wurden die mit dem „Projekt“ zusammenhängenden Daten im Informationssystem MPM analysiert. Sämtliche Daten für den Aufbau des Datenmodells mit Wissensbasis über „Projekt“ wurden selektiert, in Protégé-Ontology-Editor importiert und in OWL-Ontologien gebracht. Die Charakteristika der Properties und die Objektrestriktionen wurden per Definition ebenfalls berücksichtigt.

5.4 Integration der Datenquellen

Nachfolgend werden die zu integrierenden Datenquellen und IT-Systeme mit den gespeicherten Informationen einzeln erläutert.

Die Objekte an Bundeswasserstraßen können über das Objektkataster der WSV eindeutig identifiziert werden. Die eindeutige Objektidentifizierung erfolgt aus dem digitalen Objektkataster der WSV, der Wasserstraßen-Infrastrukturdatenbank (WInD). Die Bauteilgliederung erfolgt derzeit noch aus dem Vorgängersystem von WInD, der WADABA, welche sich derzeit in der schrittweisen Ablösung befindet [14]. Außerdem befinden sich Stammdaten in WInD, welche vom Informationssystem MPM konsumiert werden. Objektinformationen, wie z. B. über „Wehr“ und „Bundeswasserstraße“ in der Wissensbasis (s. Abbildung 3) sind in diesen IT-Systemen repräsentiert.

Zustandsinformationen der Objekte (Zustandsnoten, Teilnoten, Zustandsprognosen) des führenden Fachverfahrens WSVPruf gem. Verwaltungsvorschrift der WSV 2101 (Zielversion 2022) werden vom Informationssystem MPM konsumiert und in der Informationsdarstellung visualisiert. Objektinformationen, wie z. B. „Verschluss“, „Wehrverschluss“, „Revisionsverschluss“ und „Wehr“ in der Wissensbasis (s. Abbildung 3), sind in diesem IT-System repräsentiert.

Das Informationssystem MPM erfasst „Projekt“-Informationen, welche z. B. die Entitäten wie „Projekt“ und „Maßnahme“ (s. Abbildung 4) repräsentieren.

Hierin besteht die Beziehung, dass das Informationssystem MPM gemeinsame Daten von WInD und vom Fachverfahren WSVPruf bezieht. Geplant ist auch eine gemeinsame Datennutzung des IT-Systems Digitale Verwaltung technischer Unterlagen (DVtU). Die aktuellen Entwurf-Haushaltsunterlagen und -Ausführungsunterlagen können vom Informationssystem MPM ebenfalls konsumiert werden. Dabei können die jeweiligen Status der Unterlagen angezeigt werden.

Abbildung 5 zeigt übersichtlich diese vorgenannten Beziehungen, ohne diese explizit zu benennen.

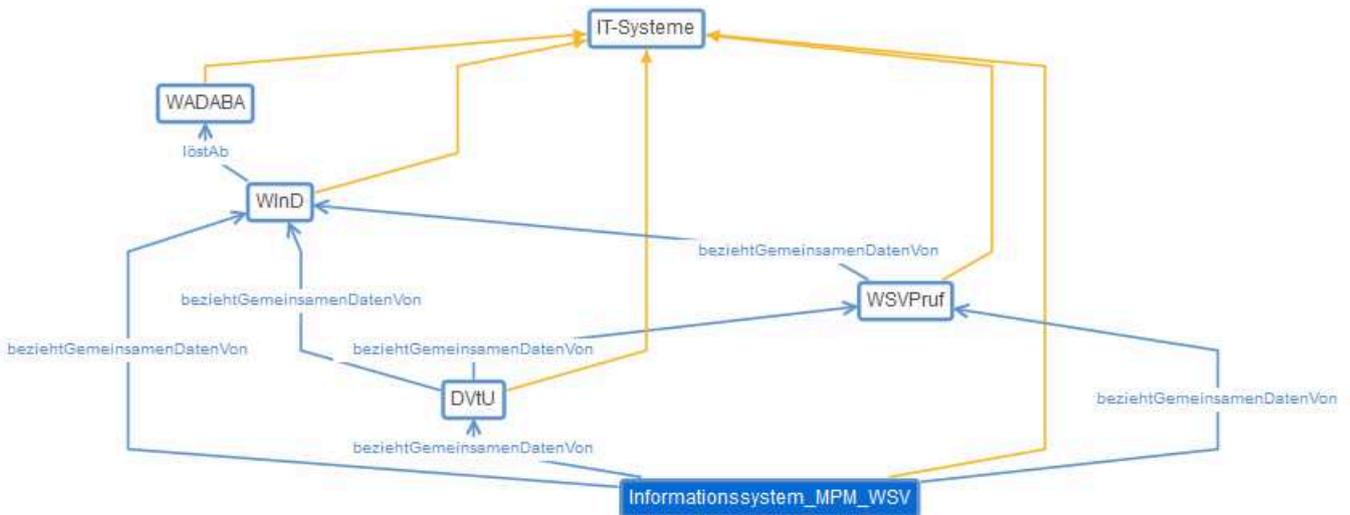


Abbildung 5: Mit dem Informationssystem MPM zusammenhängende IT-Systeme der WSV sowie die Beziehungen (modelliert in Protégé-Editor und dargestellt durch WebProtégé in <https://webprotege.stanford.edu>)

Es besteht großes Potenzial zur Vernetzung der IT-Systeme mit Unterstützung einer Wissensbasis, wenn die Daten in den Quell-Systemen von der Wissensbasis (vorgestellt in Kapitel 5.3) abgefragt werden können und der Nutzer zufriedenstellende Ergebnisse zu den skizzierten Abfragen (vorgestellt in Kapitel 5.2) vorfinden kann. Eine sukzessive Anbindung und Integration der relevanten Informationen in die Wissensbasis ist erstrebenswert.

6 Fazit und Ausblick

In diesem Beitrag wurde der erste Ansatz eines mit SWT aufgebauten Wissensbasis-gestützten Datenmodells für die Datenabfrage und -integration in die bestehenden IT-Systeme der WSV im Bereich des Verkehrswasserbaus gezeigt. Der Einsatzkontext und die damit verbundenen Anforderungen in Form von möglichen Abfragen wurden vorausgewählt. Die bisherige Arbeit fokussierte auf die Konstruktion der Wissensbasis durch eine fundierte Analyse der Begriffserklärungen im Kontext des Informationssystems MPM. Weitere Arbeit – vorgestellt in Kapitel Vorgehensweise – konzentriert sich auf die Kennzeichnung der in Ontologien befindlichen Daten mit den jeweiligen Quell-IT-Systemen, um die Vernetzungsmöglichkeiten der IT-Systeme der WSV (erläutert in Kapitel 5.4) zu skizzieren. Diese Art der Datenintegration schafft eine hohe Interoperabilität der IT-Systeme und ersetzt die derzeitigen manuellen Abfragen und manuellen Verknüpfungen der Daten in der bestehenden IT-Landschaft. Der somit erstandene Nutzen der Aufwandsreduzierung bei der Datenzusammentragung wird zusätzlich durch einen Akzeptanzfaktor bestärkt, da die bestehenden IT-Systeme, die aus den Bedürfnissen der WSV entstanden sind, erhalten bleiben.

Es kann ferner untersucht werden, inwiefern die erstellten Ontologien bei Natural Language Processing (NLP) insbesondere Information Retrieval (IR) unterstützen können. IR – eine Teilmethode aus NLP – kann effizienter Entitäten mit Hilfe von vorgegebenen Ontologien aus dem Domänenbereich identifizieren und somit den Domänenbereich mit im Vergleich zu Ontologien geringerer logischen Korrektheit abdecken. Oftmals spielt erstens die logische Korrektheit bei der semantischen Suche weniger eine Rolle als der Deckungsgrad der Domänenbereiche [6, S. 110-112]. Zweitens sind viele Wissenskomponenten in Text-Dokumenten in natürlicher Sprache und in einer unstrukturierten Form gespeichert. Aus diesen Gründen sollten die zwei vorgenannten Methoden kombiniert angewendet werden. Eine Methoden-Kombination von SWT und NLP könnte daher eine bessere Lösung für eine semantische Suchmaschine bieten.

Literatur

- [1] Informations Technik Zentrum und Bundesanstalt für Wasserbau, Hg., „Planen, Bauen und Betreiben der Wasserstraßeninfrastruktur: Empfehlungen zur Erstellung eines Masterplans Digitalisierung“, 2019. [Online] Verfügbar unter: <https://hdl.handle.net/20.500.11970/107516>. Zugriff am: 14. Juli 2021.

-
- [2] C. Heinzlmann, J. Bödefeld und Z. Duric, „Wege zur Digitalisierung im Verkehrswasserbau“ in *Dresdner Wasserbauliche Mitteilungen* 63 (2020), S. 11–20. [Online] Verfügbar unter: <https://hdl.handle.net/20.500.11970/107045>. Zugriff am: 14. Juli 2021.
- [3] *Bauwerksinformationsmodelle - Handbuch der Informationslieferungen - Teil 1: Methodik und Format, DIN EN ISO 29481-1*, 2018.
- [4] *Bauwerksinformationsmodelle - Handbuch der Informationslieferungen - Teil 2: Interaktionsframework, DIN EN ISO 29481-2*, 2017.
- [5] S. Auer, V. Bryl und S. Tramp, Hg., *Linked Open Data -- Creating Knowledge Out of Interlinked Data*. Cham: Springer International Publishing, 2014. [Online] Verfügbar unter: <https://doi.org/10.1007/978-3-319-09846-3>. Zugriff am: 14. Juli 2021.
- [6] S. Biffel und M. Sabou, Hg., *Semantic Web Technologies for Intelligent Engineering Applications*. Cham: Springer International Publishing, 2016.
- [7] P. Hitzler, M. Krötzsch, S. Rudolph und Y. Sure, *Semantic Web: Grundlagen*. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg, 2008. [Online] Verfügbar unter: <http://dx.doi.org/10.1007/978-3-540-33994-6>. Zugriff am: 14. Juli 2021.
- [8] A. Wagner, L. K. Moeller, C. Leifgen und C. Eller, *BPO: Building Product Ontology*. [Online] Verfügbar unter: <https://www.projekt-scope.de/ontologies/bpo/>. Zugriff am: 20. Juli 2021.
- [9] P. Pauwels und W. Terkaj, *ifcOWL ontology (IFC4_ADD2_TC1)*. [Online] Verfügbar unter: https://standards.buildingsmart.org/IFC/DEV/IFC4/ADD2_TC1/OWL/index.html. Zugriff am: 20. Juli 2021.
- [10] *Basic European Road OTL*. [Online] Verfügbar unter: https://www.roadotl.eu/static/eurotl-ontologies/eurotl_doc/index-en.html. Zugriff am: 20. Juli 2021.
- [11] J. Huang und R. Bergs, „Vom Objekt zum Projekt – von Daten zu Informationen“ in *Bundesanstalt für Wasserbau (Hg.): Digitalisierung im Verkehrswasserbau 2018*, S. 45–51. [Online] Verfügbar unter: <https://hdl.handle.net/20.500.11970/106114>. Zugriff am: 14. Juli 2021.
- [12] T. Huyeng, J. Huang, A. Wagner, B. Lehmann und U. Rüppel, „Digitalisierung im Verkehrswasserbau: BIM, Semantic Web und IoT-Sensorik“ in *Dresdner Wasserbauliche Mitteilungen* 62 (2019), S. 273–283. [Online] Verfügbar unter: <https://hdl.handle.net/20.500.11970/106319>. Zugriff am: 14. Juli 2021.
- [13] P. Pauwels, S. Zhang und Y.-C. Lee, „Semantic web technologies in AEC industry: A literature overview“, *Automation in Construction*, Jg. 73, November, S. 145–165, <https://doi.org/10.1016/j.autcon.2016.10.003>, 2017. Zugriff am: 14. Juli 2021.
- [14] D. Schenk und C. Laursen, „Digitalisierung im Verkehrswasserbau: Aktuelle Entwicklungen“ in *Dresdner Wasserbauliche Mitteilungen* 63 (2020), S. 89–98. [Online] Verfügbar unter: <https://hdl.handle.net/20.500.11970/107053>. Zugriff am: 14. Juli 2021.

32. Forum Bauinformatik 2021

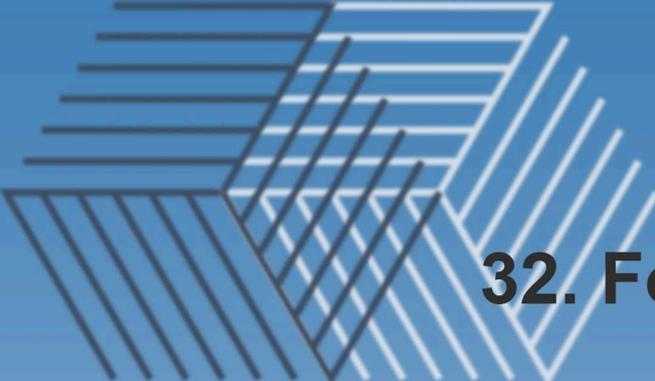
09. - 10. September 2021 in Darmstadt

Herausgeber: Michael Disser | André Hoffmann | Luisa Kuhn | Patrick Scheich



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Institut für Numerische
Methoden und Informatik
im Bauwesen



32. Forum Bauinformatik
2021

Veröffentlicht unter CC-BY 4.0 International
<https://creativecommons.org/licenses>

Quelle Tagungsband <https://doi.org/10.26083/tuprints-00019496>

