

Thema2_Modellierung_Ontologie_VersionA_Stu 22.01.26

Masterarbeit: Konzeption und semantische Modellierung eines ontologiebasierten Referenzrahmens zur Harmonisierung von Bildungsangeboten und Kompetenzanforderungen

Betreuer: Adrian Vogler, M.Sc., Prof. Dr. Matthias Hemmje

Projektbezug: HR-QDE (Human Resource Qualification Development Ecosystem)

Thema: 2 von 3 (Semantische Modellierung & Ontologie)

Notwendige Kenntnisse und Voraussetzungen

Für eine erfolgreiche Bearbeitung der Aufgabenstellung sind Vorkenntnisse in folgenden Bereichen notwendig bzw. nützlich:

Methodisch:

- Ontologien und semantische Modellierung (notwendig)
 - RDF (Resource Description Framework)
 - OWL (Web Ontology Language)
 - SPARQL
- UML (notwendig)
- Semantic Web Technologien (notwendig)
- Knowledge Representation (notwendig)
- Design Science Research (nützlich)

Technisch:

- Protégé (Ontologie-Editor) (notwendig)
 - Apache Jena Fuseki (Triple Store) (notwendig)
 - Python (nützlich)
 - RDFLib
 - ESCO API (notwendig)
 - Git (notwendig)
 - Natural Language Processing (nützlich)
 - Web-Entwicklung (nützlich)
-

Motivation und Ausgangspunkt: Der Studierendenkompass als Navigationskonzept

Studierende in grundständigen Studiengängen stehen häufig vor einem komplexen Navigationsproblem. Bildlich gesprochen gleichen sie Weltseglern, die zwar das Handwerk der Navigation beherrschen und über Seekarten (Curricula, Modulhandbücher) verfügen, deren konkreter Zielhafen (berufliche Qualifikation) jedoch oft im Nebel liegt. Während sie lernen, wie man segelt (Studium durchlaufen), bleiben die spezifischen Anforderungen der Industrie und des Arbeitsmarktes (das Ziel) oft intransparent. Es fehlt an Instrumenten, die nicht nur eine Navigation durch das Studienprogramm ermöglichen, sondern diese Navigation gezielt auf aktuell nachgefragte Berufsbilder ausrichten.

Vor diesem Hintergrund wird an verschiedenen Hochschulen das Konzept eines sogenannten „**Studierendenkompasses**“ diskutiert. Ein solches System soll als Orientierungshilfe dienen, um individuelle Lernpfade so zu gestalten, dass sie die Lücke zwischen dem akademischen Angebot und den realen Bedarfen des Arbeitsmarktes

schließen. Ziel ist es, die Navigation vom Start (Studienbeginn) bis zum Ziel (Berufsfähigkeit in einem spezifischen Feld) durch semantische Technologien und KI-gestützte Pfadempfehlungen transparent und steuerbar zu machen.

Die Betrachtung erfolgt anhand zweier beispielhafter Institutionen, die sich in ihrer Ausrichtung stark unterscheiden:

FernUniversität in Hagen (FUH): Als größte Fernuniversität Deutschlands spielt die FernUniversität eine zentrale Rolle bei der Ausbildung qualifizierter Fachkräfte, die für die Herausforderungen der Digitalisierung und Industrie 4.0 gerüstet sind. Studierende im Fernstudium benötigen eine starke Orientierungshilfe, um individuelle Lernpfade zu gestalten, da sie oft aus diversen beruflichen Kontexten kommen oder auf unterschiedliche Qualifikationsziele hinsteuern. Für diese deutschsprachigen Studierenden aus verschiedenen Ländern ist der Studierendenkompass ein Instrument zur **individuellen Orientierung** im Studium.

Hochschule Darmstadt (HDA): Die Hochschule Darmstadt ist als Teil der **European University of Technology Alliance (EUT+)** in einen europäischen Hochschulverbund eingebunden. Im Rahmen des **Bologna-Prozesses** erkennen die Partnerhochschulen von EUT+ ihre Studienleistungen gegenseitig an, was **mehrsprachige Studienprogramme** und **grenzüberschreitende Mobilität** ermöglicht. Studierende müssen verstehen, welche Kurse im Ausland ihren heimischen Anforderungen entsprechen und wie diese anerkannt werden. Hier ist der Studierendenkompass ein Instrument zur **internationalen Mobilität** und zur Sicherstellung der Vergleichbarkeit von Studienleistungen über Ländergrenzen hinweg.

Obwohl die Motivationen unterschiedlich sind (individuelle Orientierung im Fernstudium vs. internationale Mobilität im Präsenzstudium), ist die **technische Anforderung identisch**: Es wird eine **maschinenlesbare, semantische Repräsentation** von Qualifikationen und **Kompetenzen im Sinne nachgewiesener Qualifikationen** benötigt, um diese Navigation automatisiert zu unterstützen.

Gesellschaftlicher Kontext und Relevanz

Von diesem Ausgangspunkt der beiden Hochschulen aus muss der übergeordnete gesellschaftliche Kontext betrachtet werden. Obwohl die FernUniversität in Hagen und die Hochschule Darmstadt unterschiedliche Zielsetzungen verfolgen, operieren beide im gleichen sozioökonomischen Umfeld, das zunehmend durch einen gravierenden Fachkräftemangel geprägt ist (bedingt u.a. durch den demografischen Wandel).

Daraus ergibt sich für beide Institutionen die Notwendigkeit, Studierende in den grundständigen Studiengängen wesentlich gezielter und passgenauer auszubilden. Das Ziel muss sein, "brauchbare Human Resources" bereitzustellen, die dem Arbeitsmarkt ohne lange Einarbeitungszeiten zur Verfügung stehen. Um dies zu gewährleisten, gewinnen integrierte Konzepte wie das im Folgenden beschriebene Human Resource Qualification Development Ecosystem (HR-QDE) für Hochschulen massiv an Bedeutung.

Kontext: Vom Fachkräftemangel zur Qualifikationsentwicklung

Die Herausforderungen und Motivationen für die Entwicklung eines Human Resource Qualification Development Ecosystem (HR-QDE) lassen sich in zwei zentrale Perspektiven unterteilen: **Nachfrageseite (Demand-Side)** und **Angebotsseite (Supply-Side)**. Die Nachfrageseite, primär repräsentiert durch die Industrie, benötigt Fachkräfte mit spezifischen **Qualifikationen im Sinne nachgewiesener Kompetenzen**, die auf ihre Prozesse und Ressourcen zugeschnitten sind. Auf der anderen Seite können die Anbieter, wie Universitäten und Bildungseinrichtungen, durch **maschinenlesbare semantische Modelle** qualifikationsorientierte Bildungsangebote bereitstellen.

Diese klare Unterscheidung zwischen den beiden Perspektiven ermöglicht eine präzisere Adressierung von Nachfrage und Angebot. Sie dient als Grundlage für die Entwicklung eines integrierten Modells, das die Bedürfnisse der Nachfrager mit den Bildungsressourcen effektiv in Einklang bringt, um dem Fachkräftemangel nachhaltig zu begegnen.

Relevante Vorarbeiten

In diesem Kontext ordnet sich diese Arbeit in die bestehenden Forschungsaktivitäten zur Entwicklung des HR-QDE-Konzepts ein. Für diese Masterarbeit sind insbesondere folgende Vorarbeiten relevant:

FPHR (Fähigkeitzentrierte und prozessorientierte Ontologie für HR): Vogler entwickelte die FPHR-Ontologie für den HR-Kontext. Die FPHR-Ontologie verwendet die ESCO-Ontologie als obere Ontologie und modelliert **Kompetenzanforderungen** (Demand-Side). Die Integration dieser Ontologie ermöglicht die Maschinenlesbarkeit von Kompetenzanforderungen.

QBLM (Qualifications-Based Learning Model): Bedeutende Arbeiten zum QBLM wurden am Lehrstuhl von Wallenborn und Then durchgeführt. QBLM ist ein umfassendes Framework, das aus einem Domänenmodell, einem

Architekturmodell und verschiedenen Service-Verteilungsmodellen besteht. QBLM modelliert **Kursangebote und Qualifikationen** (Supply-Side) und ist ein qualifikationsbasiertes Modell zur Produktion qualifizierter Fachkräfte.

ESCO (European Skills, Competences, Qualifications and Occupations): ESCO ist ein mehrsprachiges europäisches Klassifikationssystem für Fähigkeiten, Kompetenzen, Qualifikationen und Berufe. ESCO bietet ein standardisiertes Vokabular, das als **obere Ontologie** für die Integration von Demand-Side (FPHR) und Supply-Side (QBLM) dient.

Forschungslücken und Motivation für diese Arbeit

Trotz der umfangreichen Vorarbeiten bestehen spezifische Forschungslücken im Bereich **Semantische Modellierung & Ontologie**:

1. **Fehlende Integration von FPHR und QBLM:** FPHR modelliert Demand-Side, QBLM modelliert Supply-Side, aber es fehlt ein **integriertes Modell**, das beide Seiten vereint und über ESCO als gemeinsame obere Ontologie verbindet.
 2. **Fehlende OWL-Repräsentation von QBLM:** QBLM liegt als UML-Modell vor, aber es fehlt eine **OWL-Ontologie**, die für semantische Abfragen (SPARQL) und Reasoning genutzt werden kann.
 3. **Fehlende ESCO-Annotation für Hochschulkurse:** Während ESCO primär für berufliche Weiterbildung entwickelt wurde, fehlt eine systematische Methode zur **ESCO-Annotation** von Hochschulkursen (grundständige Lehre).
-

Einordnung der Aufgabenstellung

Diese Masterarbeit leistet einen Beitrag zur Entwicklung eines übergeordneten Qualifikations-Ökosystems. Sie fokussiert sich auf die **semantische Modellierung und Integration** von Kompetenzanforderungen (Demand) und Kursangeboten (Supply) und bildet damit die konzeptionelle Grundlage für weiterführende Arbeiten zum Matching und zur Lernpfad-Generierung.

Aufgabenbeschreibung

Das Ziel dieser Masterarbeit ist die Entwicklung eines **integrierten semantischen Modells**, das:

1. **Demand-Side** (Kompetenzanforderungen) und **Supply-Side** (Kursangebote) vereint
2. **ESCO** als obere Ontologie verwendet
3. **FPHR** und **QBLM** integriert
4. Extrahierte Kompetenzen mit ESCO annotiert

Die Arbeit soll folgende Forschungsfragen beantworten:

Forschungsfragen:

2.1. Wie können extrahierte Kompetenzen mit ESCO annotiert werden?

- Welche Matching-Methoden sind geeignet (String-Matching, Semantic Matching, Machine Learning)?
- Wie kann die Qualität der Annotation bewertet werden?
- Wie können Kompetenzen, die nicht in ESCO vorhanden sind, behandelt werden?

2.2. Wie können bestehende Ontologien (ESCO, FPHR, QBLM) für HDA und FUH integriert werden?

- Wie kann QBLM von UML in OWL transformiert werden?
- Wie können FPHR (Demand-Side) und QBLM (Supply-Side) über ESCO integriert werden?
- Wie können HDA-spezifische und FUH-spezifische Konzepte modelliert werden?

2.3. Wie können Kompetenzanforderungen (Demand) und Kursangebote (Supply) in einem einheitlichen maschinenlesbaren semantischen Modell repräsentiert werden? Dabei sollen die Felder "Inhalt" und "Ziel" konkateniert werden, um eine umfassendere semantische Analyse zu ermöglichen.

- Welche Konzepte müssen modelliert werden (Courses, Competencies, Job Postings, Learning Outcomes)?
 - Welche Relationen müssen modelliert werden (hasCompetency, requiresCompetency, hasPrerequisite)?
 - Wie kann das Modell in RDF/OWL repräsentiert werden?
-

Phasen der Arbeit

Die Arbeit gliedert sich in vier Phasen nach der Forschungsmethodik von Nunamaker:

Phase 1: Beobachtung (Analyse und Recherche)

Ziel: Verstehen der Ontologien und Annotationsmethoden

Aktivitäten:

- Literaturrecherche zu Ontologie-Engineering, Semantic Web, ESCO, Ontologie-Alignment
- Analyse der Vorarbeiten (FPHR, QBLM, ESCO)
- Analyse der Eingangsdaten (extrahierte Kompetenzen)

Ergebnis:

- Literaturübersicht zu Ontologie-Engineering und ESCO
 - Dokumentation der Vorarbeiten
 - Analyse der Eingangsdaten
-

Phase 2: Theoriebildung (Konzeption und Modellierung)

Ziel: Konzeption des integrierten semantischen Modells

Aktivitäten:

- Konzeption der ESCO-Annotation (Matching-Methode, Qualitätsbewertung)
- Konzeption der Ontologie-Integration (Transformation QBLM → OWL, Alignment FPHR + QBLM)
- Modellierung von HDA/FUH-spezifischen Konzepten
- Definition der Konzepte und Relationen

Ergebnis:

- Konzept für ESCO-Annotation
 - Architekturdiagramm der Ontologie-Integration
 - UML/OWL-Modell des integrierten semantischen Modells
-

Phase 3: Systementwicklung (Implementierung)

Ziel: Implementierung des integrierten semantischen Modells

Aktivitäten:

- Implementierung der ESCO-Annotation (Python-Tool, ESCO API Integration)
- Transformation von QBLM in OWL (Protégé)
- Integration von FPHR und QBLM
- Aufbau des RDF Triple Stores (Apache Jena Fuseki)

Ergebnis:

- ESCO-Annotationstool (Python)
 - QBLM-OWL-Ontologie
 - Integrierte HR-QDE-Ontologie (FPHR + QBLM + ESCO)
 - RDF Triple Store mit annotierten Daten
-

Phase 4: Experimentierung (Evaluation)

Ziel: Evaluation der Annotationsqualität und Modellkonsistenz

Aktivitäten:

- Evaluation der ESCO-Annotation (Precision, Recall, F1-Score)

- Evaluation der Ontologie (Konsistenzprüfung, Vollständigkeit, Korrektheit)
- SPARQL-Abfragen zur Validierung
- Identifikation von Verbesserungspotenzialen

Ergebnis:

- Evaluationsbericht mit Precision, Recall, F1-Score
 - Konsistenzbericht der Ontologie
 - Dokumentation von Beispiel-SPARQL-Abfragen
 - Empfehlungen für Verbesserungen
-

Erwartete Ergebnisse

Die Arbeit soll folgende konkrete Ergebnisse liefern:

1. **ESCO-Annotationstool:** Python-basiertes Tool zur automatischen ESCO-Annotation von Kompetenzen
 2. **Das Referenzmodell (OWL):**
 - QBLM-OWL-Ontologie (Transformation von UML)
 - Integrierte HR-QDE-Ontologie (FPHR + QBLM + ESCO)
 3. **Workflow-Modelle:** Dokumentation der Annotierungs- und Integrationsprozesse
 4. **RDF Triple Store:** Apache Jena Fuseki mit integrierter Ontologie und annotierten Daten
 5. **Evaluationsbericht:** Dokumentation der Annotationsqualität und Modellkonsistenz
 6. **Wissenschaftliche Dokumentation:** Abschlussbericht (Masterarbeit)
-

Zeitplan

Die Arbeit soll innerhalb von **6 Monaten** abgeschlossen werden:

- **Monat 1:** Phase 1 (Beobachtung)
 - **Monat 2:** Phase 2 (Theoriebildung)
 - **Monat 3-4:** Phase 3 (Systementwicklung)
 - **Monat 5:** Phase 4 (Experimentierung)
 - **Monat 6:** Abschlussdokumentation
-

Literatur

Wichtige Referenzen:

- Vogler, A. (2022). Fähigkeitenzentrierte und prozessorientierte Ontologie für den HR-Kontext (FPHR). Masterarbeit, FernUniversität Hagen.
- Wallenborn, B. & Then, M. (2018-2020). Qualifications-Based Learning Model (QBLM). FernUniversität Hagen.
- ESCO (European Skills, Competences, Qualifications and Occupations). <https://esco.ec.europa.eu/>
- Hitzler, P., Krötzsch, M., & Rudolph, S. (2009). Foundations of Semantic Web Technologies. CRC Press.
- Gruber, T. R. (1993). A translation approach to portable ontology specifications. Knowledge Acquisition, 5(2), 199-220.