

# Thema3\_Matching\_System\_VersionA\_Studierende 22.01.26

## Masterarbeit: Entwicklung eines KI-gestützten Matching-Systems und einer Mediator-Architektur zur Generierung personalisierter Lernpfade

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

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

**Thema:** 3 von 3 (Matching-Algorithmus & Systemarchitektur)

---

### Notwendige Kenntnisse und Voraussetzungen

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

#### Methodisch:

- Software-Architekturen (notwendig)
  - Service-orientierte Architekturen
  - Mediator-Wrapper-Architekturen
  - Verteilte Systeme
- AI Planning (notwendig)
- Graph Algorithms (notwendig)
- Recommender Systems (nützlich)
- User-Centered Design (nützlich)
- Design Science Research (nützlich)

#### Technisch:

- Python (notwendig)
  - Graph-Bibliotheken (notwendig)
    - NetworkX
    - igraph
  - Web-Entwicklung (notwendig)
    - Flask oder FastAPI (Backend)
    - React oder Vue.js (Frontend)
  - SPARQL (notwendig)
  - Semantic Web Technologien (nützlich)
  - Docker (nützlich)
  - Git (notwendig)
- 

### 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:

**SQW (Semantic Qualification Web):** Die Dissertation von Lothary zum SQW beschreibt ein System, das auf Basis semantischer Technologien und KI die Generierung personalisierter, lebenslanger Lernpfade unterstützt. Das zentrale

Element von SQW ist der **Lernpfad-Generierungsalgorithmus**, der Graph-basiertes Pathfinding verwendet (Dijkstra, A\*), Kompetenzlücken zwischen aktuellem Profil und Zielprofil berechnet, Kurse findet, die die Lücken schließen, und Voraussetzungen sowie ECTS berücksichtigt.

**KM-EP (Knowledge Management Ecosystem Portal):** Das KM-EP wurde als Plattform verwendet, da es einfache und komfortable Autorenwerkzeuge für Kurse bot und aufgrund seiner modularen Struktur gut geeignet war, offene Herausforderungen anzugehen. Es spielte eine Schlüsselrolle bei der Bereitstellung von Werkzeugen und Unterstützung für Autoren und Studierende zur Erstellung und Verwaltung qualifikationsbasierter Kurse und Programme.

---

## Forschungslücken und Motivation für diese Arbeit

Trotz der umfangreichen Vorarbeiten bestehen spezifische Forschungslücken im Bereich **Matching & Systemarchitektur**:

1. **Fehlende institutionsübergreifende Lernpfad-Generierung:** SQW fokussiert auf eine einzelne Institution. Es fehlt ein Algorithmus, der **institutionsübergreifende Lernpfade** (HDA ↔ FUH) generiert und dabei unterschiedliche Studienformen (Präsenz vs. Fern) berücksichtigt.
2. **Fehlende Integration von Demand-Side und Supply-Side:** SQW fokussiert primär auf Supply-Side (Kursangebote). Es fehlt eine systematische Methode, um **Kompetenzanforderungen** (Demand-Side) mit **Kursangeboten** (Supply-Side) abzugleichen.
3. **Fehlende Benutzeroberflächen für Studierende:** Es fehlen intuitive Benutzeroberflächen, die Studierenden ermöglichen, ihr Profil zu erfassen, Ziele zu setzen und personalisierte Lernpfade zu erhalten.

---

## Einordnung der Aufgabenstellung

Diese Masterarbeit leistet einen Beitrag zur Entwicklung eines übergeordneten Qualifikations-Ökosystems. Sie fokussiert sich auf die **Entwicklung von Matching-Algorithmen und Systemarchitektur** für die Generierung personalisierter Lernpfade und bildet damit die technische Umsetzung des Studierendenkompasses.

---

## Aufgabenbeschreibung

Das Ziel dieser Masterarbeit ist die Entwicklung eines **KI-gestützten Systems zur Generierung personalisierter Lernpfade**, das:

1. **Kompetenzlücken** zwischen aktuellem Profil und Zielprofil berechnet
2. **Personalisierte Lernpfade** generiert, die die Lücken schließen
3. **Institutionsübergreifende Kursempfehlungen** (HDA ↔ FUH) erstellt
4. **Technisch den Abgleich** zwischen Kompetenzanforderungen (Demand) und Kursangeboten (Supply) ermöglicht

Die Arbeit soll folgende Forschungsfragen beantworten:

### Forschungsfragen:

**3.1. Wie können personalisierte Lernpfade basierend auf Kompetenzlücken generiert werden, wobei eine prozentuale Übereinstimmung zwischen akademischen Modulen und Weiterbildungsangeboten berechnet wird?**

- Wie können Kompetenzlücken berechnet werden ( $\text{Gap} = \text{Zielprofil} - \text{Aktuelles Profil}$ )?
- Welche Algorithmen sind geeignet (Graph-basiertes Pathfinding, AI Planning)?
- Wie können Voraussetzungen, ECTS, Zeitbudget berücksichtigt werden?

**3.2. Wie können institutionsübergreifende Kursempfehlungen (HDA ↔ FUH) erstellt werden?**

- Wie können Kurse von HDA und FUH verglichen werden (ESCO-Annotation)?
- Wie können unterschiedliche Studienformen berücksichtigt werden (Präsenz vs. Fern)?
- Wie können Anerkennungsregeln modelliert werden (Bologna-Prozess)?

**3.3. Wie kann das System technisch den Abgleich zwischen Kompetenzanforderungen und Kursangeboten ermöglichen? Welche Rolle spielen Embedding-basierte Vektorraum-Vergleiche (z.B. OpenAI oder HuggingFace Embeddings) für semantische Ähnlichkeitsberechnungen?**

- Wie können Stellenanzeigen (Demand-Side) mit Kursen (Supply-Side) gematcht werden?
  - Wie kann die Qualität des Matchings bewertet werden?
  - Wie können Lücken im Curriculum identifiziert 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 Algorithmen und Benutzeranforderungen

**Aktivitäten:**

- Literaturrecherche zu AI Planning, Graph Algorithms, Recommender Systems
- Analyse der Vorarbeiten (SQW, KM-EP)
- Analyse der Eingangsdaten (semantisches Modell, RDF Triple Store)
- Anforderungsanalyse (Studierende, Lehrende)

**Ergebnis:**

- Literaturübersicht zu AI Planning und Graph Algorithms
  - Dokumentation der Vorarbeiten
  - Analyse der Eingangsdaten
  - Anforderungskatalog
- 

### Phase 2: Theoriebildung (Konzeption und Modellierung)

**Ziel:** Konzeption der Algorithmen und Benutzeroberflächen

**Aktivitäten:**

- Konzeption des Lernpfad-Algorithmus (Gap-Berechnung, Pathfinding, Constraints)
- Konzeption des Matching-Algorithmus (Demand-Supply-Abgleich, Qualitätsbewertung)
- Konzeption der Benutzeroberflächen (Studierende, Lehrende)
- Konzeption der Systemarchitektur (Mediator-Wrapper, Backend, Frontend)

**Ergebnis:**

- Algorithmus-Spezifikation für Lernpfad-Generierung
  - Algorithmus-Spezifikation für Matching
  - Wireframes und Mockups für Benutzeroberflächen
  - Architektordiagramm des Gesamtsystems
- 

### Phase 3: Systementwicklung (Implementierung)

**Ziel:** Implementierung des Studierendenkompasses

**Aktivitäten:**

- Implementierung des Lernpfad-Algorithmus (Python, NetworkX, SPARQL)
- Implementierung des Matching-Algorithmus (Python, SPARQL)
- Implementierung der Benutzeroberflächen (Backend: Flask/FastAPI, Frontend: React/Vue.js)
- Integration mit KM-EP
- Integration mit HIS in One (HDA) und Moodle (FUH), soweit APIs verfügbar

**Ergebnis:**

- Lernpfad-Generierungsalgorithmus (Python)

- Matching-Algorithmus (Python)
  - Web-basierte Benutzeroberflächen (Backend + Frontend)
  - Deployment auf KM-EP
- 

## Phase 4: Experimentierung (Evaluation)

**Ziel:** Evaluation der Lernpfade und Benutzeroberflächen

**Aktivitäten:**

- Evaluation des Lernpfad-Algorithmus (Beispiel-Lernpfade, Expertenbewertung, Laufzeit)
- Evaluation des Matching-Algorithmus (Precision, Recall, Coverage)
- Usability-Evaluation (Cognitive Walkthrough, System Usability Scale)
- Identifikation von Verbesserungspotenzialen

**Ergebnis:**

- Evaluationsbericht für Lernpfad-Algorithmus
  - Evaluationsbericht für Matching-Algorithmus
  - Usability-Evaluationsbericht
  - Empfehlungen für Verbesserungen
- 

## Erwartete Ergebnisse

Die Arbeit soll folgende konkrete Ergebnisse liefern:

1. **Lernpfad-Generierungsalgorithmus:** Python-basierter Algorithmus zur Generierung personalisierter Lernpfade
  2. **Matching-Algorithmus:** Python-basierter Algorithmus zum Abgleich zwischen Demand und Supply
  3. **Systemarchitektur und integrierter Prototyp:**
    - Mediator-Wrapper-Architektur für HDA und FUH
    - Web-basierte Benutzeroberfläche für Studierende
    - Web-basierte Benutzeroberfläche für Lehrende
    - Integration mit RDF Triple Store (SPARQL)
    - Deployment auf KM-EP
  4. **Beispiel-Lernpfade:** Generierte Lernpfade für verschiedene Zielprofile
  5. **Evaluationsbericht:** Dokumentation der Evaluation
  6. **Wissenschaftliche Dokumentation:** Abschlussbericht (Masterarbeit)
- 

## Ausblick und Future Work

Die vorliegende Arbeit fokussiert sich auf die **technische Machbarkeit** des Abgleichs zwischen Kompetenzanforderungen und Kursangeboten. Eine weiterführende Fragestellung, die über den Rahmen dieser Arbeit hinausgeht, wäre die **qualitative Bewertung**, ob das Kursangebot einer Hochschule die Anforderungen der Industrie tatsächlich deckt (Gap-Analyse auf Inhaltsebene). Diese Fragestellung ist politisch sensibel und würde eine umfassende curriculare Analyse erfordern, die in zukünftigen Arbeiten adressiert werden könnte.

---

## 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:

- Lothary, S. (2024). Semantic Qualification Web (SQW). Dissertation, FernUniversität Hagen.
- Russell, S., & Norvig, P. (2020). Artificial Intelligence: A Modern Approach (4th ed.). Pearson.
- Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). Introduction to Algorithms (3rd ed.). MIT Press.
- Ricci, F., Rokach, L., & Shapira, B. (2015). Recommender Systems Handbook (2nd ed.). Springer.
- NetworkX Documentation. <https://networkx.org/>