

Exposé für:

**Konzeption und Architekturmodell zur MCP-Integration in CMS anhand
von TYPO3 für KI-kompatible Informationsbereitstellung im öffentlichen
Sektor**

Bachelor-Arbeit

zur Erlangung des Grades Bachelor of Science in Medieninformatik
des Fachbereichs Informatik an der
Technischen Hochschule Brandenburg

vorgelegt von:

Philip Seibold

Betreuer: Sebastian Kreideweiß, M.Sc.

Zweitgutachter: Prof. Dr. rer. nat. Martin Christof Kindsmüller

Seibold, Philip Bachelor of Science Studiengang Medieninformatik	Start: Ende: 30.09.2025
--	----------------------------

Jena den 01.06.2025

Inhaltsverzeichnis

Abbildungsverzeichnis	III
Tabellenverzeichnis	IV
Abkürzungsverzeichnis	V
1 Problemstellung	1
2 Erkenntnisinteresse	3
3 Fragestellung	4
4 Ziele und Hypothesen	5
5 Theoriebezug	7
6 Forschungsstand	8
7 Methode	9
7.1 Literatur- und Dokumentenanalyse.....	9
7.2 Qualitative Interviews	9
7.3 Fragebogenerstellung und Durchführung.....	10
7.4 Wissenschaftlicher Anspruch	10
8 Material	11
9 Gliederungsentwurf	12
10 Vorläufiges Literaturverzeichnis	13
11 Grober Zeitplan	14

Abbildungsverzeichnis

Es konnten keine Einträge für ein Abbildungsverzeichnis gefunden werden.

Tabellenverzeichnis

Es konnten keine Einträge für ein Abbildungsverzeichnis gefunden werden.

Abkürzungsverzeichnis

CMS	Content-Management-Systeme
HS	Hochschule
THB	Technische Hochschule Brandenburg

1 Problemstellung

In einer zunehmend digitalisierten Verwaltung und öffentlichen Informationsbereitstellung wächst der Bedarf an intelligenten Systemen, die nicht nur Daten speichern, sondern auch kontextualisiert zugänglich machen. Klassische Content-Management-Systeme (CMS) wie TYPO3 stoßen dabei an Grenzen, da Inhalte primär zur menschlichen Rezeption strukturiert werden. Gleichzeitig gewinnen KI-gestützte Verfahren wie Retrieval-Augmented Generation (RAG) an Bedeutung, die externe Wissensquellen zur Verbesserung von Antworten dynamisch einbeziehen. Damit solche Systeme zuverlässig funktionieren, benötigen sie Zugriff auf strukturierte, semantisch angereicherte Inhalte. Es ist derzeit davon auszugehen, dass dieser Trend sich weiter verstärken wird:

Laut einer repräsentativen Studie von Bitkom¹ (2025) haben bereits 67 % der deutschen Bevölkerung generative KI-Anwendungen wie ChatGPT oder Gemini genutzt die Tendenz ist dabei steigend. KI-Systeme ersetzen zunehmend klassische Recherchewege über Suchmaschinen und Webportale.

Medizinische Studien zeigen Halluzinationsraten von bis zu 39,6 % bei GPT-3.5, 28,6 % bei GPT-4 und in spezifischen Fällen bis zu 91,4 % bei Gemini (ehemals Bard)².

Sie reichen von ungenauen Daten bis hin zu komplett falschen Antworten. Für den öffentlichen Sektor ist das ein Risiko: Bürger*innen verlassen sich auf korrekte, aktuelle und nachvollziehbare Inhalte. Werden diese über KI-Systeme fehlerhaft transportiert, drohen Vertrauensverlust, Fehlentscheidungen oder gar Desinformation. Es ergibt sich somit eine dringende Notwendigkeit, öffentlich bereitgestellte Inhalte systematisch KI-kompatibel aufzubereiten.

Während das Model Context Protocol (MCP) im Drupal-Ökosystem bereits erste Wege zur strukturierten Kontextbereitstellung aufzeigt, fehlt es aktuell für TYPO3 an vergleichbaren Ansätzen. Die Integration semantischer Informationsmodelle und standardisierter Schnittstellen zur KI-Nutzung steht

¹ https://www.bitkom.org/Presse/Presseinformation/KI-Nutzung-boomt-Angst-vor-Abhaengigkeit-Ausland-gross?utm_source=chatgpt.com#

² <https://pubmed.ncbi.nlm.nih.gov/38776130/>

hier noch am Anfang. Daraus ergibt sich eine Forschungslücke, die diese Arbeit adressieren möchte.

Das CMS TYPO3 wurde aufgrund der aktuell begonnen Migration³ von 90+ Bedarfsträgern auf Bundesebene in Deutschland mit über 500 Websites hin zur bundeseigenen CMS-Lösung GovernmentSiteBuilder (kurz GSB) Version 11 auf Basis TYPO3 (Version 12) ausgewählt. Eine Optimierung des Zusammenspiels der Inhalte aus dem CMS in anfragende KI-Lösungen oder umgekehrt kann eine signifikante Optimierung im Bereich der Verwaltungsdigitalisierung bedeuten.

³ <https://www.materna.de/newshub/presse/digitale-souveraenitaet-staerken-materna-und-leistungstarkes-partnernetzwerk-gewinnen-rahmenvertrag>

2 Erkenntnisinteresse

Die zentrale Motivation dieser Arbeit besteht darin zu untersuchen, wie sich das Model Context Protocol (MCP) konzeptionell in TYPO3 integrieren lässt, um die Inhalte des CMS insbesondere im öffentlichen Kontext für moderne KI-Systeme besser nutzbar zu machen. Dabei interessiert insbesondere, welche strukturellen und semantischen Voraussetzungen erfüllt sein müssen, um Inhalte so aufzubereiten, dass sie maschinenlesbar, kontextsensitiv und wiederverwendbar werden. Die Arbeit verfolgt keinen prototypischen Implementierungsansatz, sondern zielt auf die Erstellung eines belastbaren Konzepts, das als Grundlage für künftige Entwicklungen dienen kann. Im Vordergrund stehen die Adaptierbarkeit bestehender Lösungen (z. B. aus dem Drupal-Ökosystem), die Rolle semantischer Datenformate (z. B. JSON-LD, Schema.org) sowie die Anwendungsrelevanz für öffentliche Institutionen.

3 Fragestellung

Wie kann das Model Context Protocol in einem TYPO3-basierten CMS konzeptionell integriert werden, um eine standardisierte, strukturierte und KI-fähige Kontextbereitstellung von Inhalten zu ermöglichen?

4 Ziele und Hypothesen

Ziele der Arbeit:

- Darstellung der Funktionsweise und des Mehrwerts von Retrieval-Augmented Generation (RAG) im Kontext öffentlich bereitgestellter Informationen
- Analyse der konzeptionellen Grundlagen des MCP im Drupal-Kontext
- Übertragbarkeit und mögliche Anpassungen von MCP für TYPO3 herausarbeiten
- Entwicklung eines konzeptionellen Modells zur MCP-basierten Kontextbereitstellung in TYPO3
- Ableitung konkreter Anforderungen an Datenstrukturen, Schnittstellen und Redaktionsprozesses
- Reflexion der Eignung solcher Konzepte für typische Szenarien im öffentlichen Sektor (z. B. Klinikportale, Ministeriumswebseiten)

Hypothesen:

- Eine konzeptionelle Adaption des Model Context Protocols (MCP) in TYPO3 verbessert die technische Anschlussfähigkeit an Retrieval-Augmented-Generation-(RAG)-Systeme. was sich anhand eines entwickelten Architekturmodells und dem Vergleich mit bestehenden CMS-Lösungen (z. B. Drupal) nachvollziehbar darstellen lässt.
- Die Nutzung von JSON-LD und Schema.org im Kontext öffentlicher Websites wird von Stakeholdern als verständlich, nachhaltig und interoperabel bewertet.
- Öffentliche Institutionen haben einen hohen Bedarf an nachvollziehbarer, autoritativer KI-Kommunikation, der durch eine MCP-basierte Kontextbereitstellung wirksam unterstützt werden kann.

Erfolgskriterien der Arbeit:

Die Bachelorarbeit ist erfolgreich, wenn auf Basis der theoretischen und empirischen Erkenntnisse ein konsistentes, nachvollziehbares Konzept zur Integration des Model Context Protocols (MCP) in TYPO3 vorliegt. Dieses Konzept muss:

- die konzeptionelle Übertragung der MCP-Idee aus dem Drupal-Ökosystem auf TYPO3 leisten,
- eine Systemarchitektur und mögliche Schnittstellen zur semantischen Anreicherung beschreiben,
- für den öffentlichen Sektor typische Anwendungsfälle abbilden (z. B. Behördenportale),

- auf Praxiserkenntnissen aus Interviews fundieren,
- und den Transfer des Konzepts in CMS-Infrastrukturen nachvollziehbar beschreiben.

Ein prototypisches System oder eine Implementierung ist nicht vorgesehen. Die Validität der Ergebnisse ergibt sich aus der methodisch abgesicherten Konzeptentwicklung und der Einordnung durch Expert*innen.

5 Theoriebezug

Die vorliegende Arbeit bewegt sich an der Schnittstelle von Content-Management-Systemen (CMS), semantischer Informationsmodellierung und Künstlicher Intelligenz (KI), insbesondere im Kontext von Retrieval-Augmented Generation (RAG). Theoretisch basiert sie auf drei zentralen Bezugspunkten:

Erstens stellt die Architektur von Content-Management-Systemen (wie TYPO3 oder Drupal) den Rahmen für die Erzeugung und Bereitstellung von Inhalten dar. Während klassische CMS auf redaktionelle Nutzung und HTML-basierte Ausspielung abzielen, rückt mit dem Aufkommen intelligenter Systeme die maschinelle Lesbarkeit in den Fokus.

Zweitens greifen semantische Webtechnologien wie JSON-LD (JavaScript Object Notation for Linked Data) und Schema.org grundlegende Ideen der formalen Bedeutungsbeschreibung auf. Diese Formate erlauben es, Inhalte nicht nur zu strukturieren, sondern auch ihre Bedeutung explizit auszuzeichnen, wodurch automatisierte Systeme den Kontext erfassen und interpretieren können.

Drittens basiert die Anbindung an KI-Systeme auf dem Konzept von Retrieval-Augmented Generation (RAG). RAG verbindet Large Language Models (LLMs) mit einer Retrieval-Komponente, die es ermöglicht, domänenspezifische, aktuelle oder autoritative Daten in die Antwortgenerierung einzubeziehen. Damit RAG effektiv funktionieren kann, müssen Inhalte standardisiert, strukturiert und kontextualisiert vorliegen.

Die Arbeit stützt sich somit theoretisch auf Konzepte aus der semantischen Webforschung, der KI-Architektur (insbesondere LLMs mit Retrieval-Komponente) und der Informationsstrukturierung in CMS. Im Zentrum steht die Frage, wie sich diese Bereiche in ein konsistentes Modell überführen lassen, das in TYPO3 realisierbar ist.

6 Forschungsstand

Die Forschung zur Integration semantischer Strukturen in CMS und deren Nutzbarmachung für KI-Systeme ist aktuell noch fragmentarisch. Einzelne Beiträge beschäftigen sich mit der Anreicherung von Inhalten durch strukturierte Metadaten (z. B. mittels JSON-LD oder schema.org), doch systematische Studien zu deren Nutzung im Kontext von Retrieval-Augmented Generation (RAG) fehlen weitgehend. Ein innovativer Impuls kommt aus der Drupal-Community, in der das Model Context Protocol (MCP) 2024 vorgestellt wurde. Ziel des MCP ist es, über standardisierte Schnittstellen Inhalte aus CMS für KI-Anwendungen besser zugänglich zu machen, indem sie in deklarative, modulare Kontexte gegliedert und maschinenlesbar gemacht werden. Erste Implementierungen für Drupal liegen bereits vor (vgl. drupalmcp.io), doch vergleichbare Ansätze für TYPO3 fehlen bislang.

Die Diskussion in Praxis und Politik – insbesondere im öffentlichen Sektor – gewinnt ebenfalls an Fahrt. Im Rahmen von Projekten wie KIPITZ (Kooperations-Initiative Public Interest Tech Zentrum), betrieben durch das ITZBund, wird diskutiert, wie Behörden und Ministerien digitale Infrastrukturen aufbauen können, die für KI-Anwendungen geeignet sind, ohne die Informationshoheit aufzugeben. Auch der wissenschaftliche Diskurs um transparente KI-Systeme (Golem, 2024; Horstmann, 2024) betont die Relevanz verlässlicher und nachvollziehbarer Datenquellen im Dialog mit Nutzenden. Für TYPO3, das in vielen öffentlichen Einrichtungen im Einsatz ist, besteht hier ein erheblicher Entwicklungsbedarf – und zugleich eine Chance: Durch konzeptionelle Erweiterungen wie das MCP könnten Inhalte aus TYPO3 systematisch erschlossen und KI-kompatibel bereitgestellt werden. Diese Forschungslücke greift die Arbeit auf und versucht, mit einer konzeptionellen Analyse und einem Übertragungsmodell einen Beitrag zur Diskussion zu leisten.

7 Methode

Da es sich bei dieser Bachelorarbeit um eine konzeptionell ausgerichtete Arbeit handelt, steht nicht die technische Implementierung, sondern die systematische Entwicklung eines Übertragungsmodells für das Model Context Protocol (MCP) in TYPO3 im Vordergrund. Methodisch wird ein qualitativ-analytischer Zugang gewählt, der aus einer Kombination von strukturierter Literatur- und Dokumentenanalyse sowie einer qualitativen Expertenbefragung besteht.

7.1 Literatur- und Dokumentenanalyse

Zur Klärung der theoretischen und technischen Grundlagen erfolgt zunächst eine strukturierte Analyse von Dokumentationen, offiziellen Spezifikationen (z. B. JSON-LD, Schema.org, MCP-API), Open-Source-Implementierungen (z. B. aus Drupal), Fachliteratur und relevanten Beiträgen aus der Praxis (z. B. Golem-Artikel, KIPITZ-Initiative). Diese dient der Beschreibung des Problemraums, der Einordnung bestehender Lösungen und der Entwicklung eines Kategoriensystems für die Konzeptentwicklung.

7.2 Qualitative Interviews

Zur Einordnung der Relevanz und Realisierbarkeit des MCP in TYPO3 sowie zur Perspektivensammlung aus der Praxis werden qualitative Leitfadeninterviews mit ausgewählten Expertinnen durchgeführt. Ziel ist es, unterschiedliche Blickwinkel auf Anforderungen, Anwendungsfelder und technische Machbarkeit zu erfassen. Dabei kommen Gesprächspartnerinnen aus folgenden Bereichen in Betracht:

- Entwicklerinnen und Integratorinnen aus dem TYPO3-Umfeld
- Expert*innen aus der Drupal-MCP-Community
- IT-Dienstleister und Projektverantwortliche im öffentlichen Sektor
- Stakeholder mit KI-Strategiebezug (z. B. im Kontext von KIPITZ)

7.3 Fragebogenerstellung und Durchführung

Für die Interviews wird ein standardisierter Fragenkatalog entwickelt, der auf offenen Fragen basiert und explorativ angelegt ist. Er soll einen flexiblen Gesprächsverlauf ermöglichen, gleichzeitig aber eine systematische Vergleichbarkeit zwischen den Interviewpartnern fördern. Der Fragenkatalog berücksichtigt sowohl technische als auch strategische Aspekte, wie:

- Bewertung des MCP im Hinblick auf Interoperabilität
- Einschätzung der Relevanz semantischer Modelle (z. B. JSON-LD)
- Praktische Herausforderungen der Integration in TYPO3
- Potenziale und Risiken im öffentlichen Kontext (z. B. mit Blick auf Informationssouveränität)

Die Interviews werden (sofern zugestimmt) dokumentiert und in Form qualitativer Stichwortanalysen ausgewertet. Die Ergebnisse fließen in die konzeptionelle Ausarbeitung ein, insbesondere im Kapitel zur Markt- und Bedarfsanalyse sowie zur Ableitung technischer Komponenten des Modells.

7.4 Wissenschaftlicher Anspruch

Die methodische Anlage orientiert sich an den Standards qualitativer Sozialforschung (vgl. Lamnek, Flick) und zielt darauf ab, Erkenntnisse aus Expert*innenwissen systematisch in die Entwicklung eines praxisorientierten, zugleich generalisierbaren Konzepts zu überführen. Die gewonnenen Aussagen dienen nicht der statistischen Verallgemeinerung, sondern der fundierten Reflexion über Bedarfe, Strukturen und Handlungsspielräume im Umgang mit KI-kompatiblen CMS-Systemen.

Die Arbeit verfolgt einen konzeptionellen Ansatz. Es werden qualitative Methoden der Literatur- und Systemanalyse genutzt, ergänzt durch einen strukturierenden Vergleich zwischen Drupal MCP und TYPO3. Darauf aufbauend wird ein konzeptionelles Architekturmodell entwickelt. Eine beispielhafte Validierung durch Mapping auf reale CMS-Beispiele (zum Beispiel Seitenstruktur des Bundesgesundheitsministeriums) ist optional geplant.

8 Material

- Dokumentation und Quellcode des MCP-Projekts (GitHub, drupalmcp.io)
- TYPO3-Dokumentation, insbesondere zur Extension-Entwicklung und zu Content-Strukturen
- Datenstrukturen und beispielhafte Inhalte öffentlicher Websites (zum Beispiel bundesgesundheitsministerium.de, bundes-klinik-atlas.de)
- Fachliteratur zu semantischer Kontextmodellierung, RAG, JSON-LD und LLM-Architektur

9 Gliederungsentwurf

1. Einleitung

2. Hintergrund & Theoretische Grundlagen

- 2.1 TYPO3 und CMS im KI-Kontext
- 2.2 Semantische Datenformate
- 2.3 LLMs und Retrieval-Augmented Generation (RAG)
- 2.4 Model Context Protocol (MCP)
- 2.5 Methodischer Rahmen der Konzeptentwicklung

3. Marktanalyse und Bedarfsermittlung

- 3.1 Relevanz von MCP für den CMS-Markt
- 3.2 Vergleich existierender Systeme
- 3.3 Stakeholder- und Zielgruppenanalyse

4. Konzeptentwicklung zur MCP-Integration in TYPO3

- 4.1 Anforderungen und Use Cases
- 4.2 Systemarchitektur
- 4.3 Schnittstellenbeschreibung und Datenfluss
- 4.4 Umsetzungsszenario/Implementierungsszenario

5. Analyse & Bewertung

- 5.1 Zielerfüllung und Anforderungskontrolle
- 5.2 Vergleich mit anderen Ansätzen
- 5.3 Chancen und Herausforderungen

6. Ausblick

- 6.1 Weiterentwicklungsmöglichkeiten
- 6.2 Einsatzpotenzial in CMS und KI-Systemen

10 Vorläufiges Literaturverzeichnis

...

11 Grober Zeitplan

Folgender Zeitplan besteht.

Kalenderwoche	Zeitraum	Aufgabe
KW 27	30.06.2025 – 06.07.2025	Einleitung, Themenabgrenzung und Zielsetzung (Kapitel 1)
KW 28	07.07.2025 – 13.07.2025	Grundlagen: CMS, semantische Daten, LLMs, RAG, MCP, Methode (Kapitel 2)
KW 29	14.07.2025 – 20.07.2025	Fortführung Grundlagen, Methodik auf Konzeptarbeit zuschneiden (Kapitel 2.5)
KW 30	21.07.2025 – 27.07.2025	Marktanalyse: MCP-Relevanz, Vergleich Systeme, Stakeholder (Kapitel 3)
KW 31	28.07.2025 – 03.08.2025	Interview- und Fragebogenphase zur Bedarfsermittlung (Kapitel 3.3)
KW 32	04.08.2025 – 10.08.2025	Konzeptentwicklung: Anforderungen, Use Cases, Architektur (Kapitel 4.1–4.2)
KW 33	11.08.2025 – 17.08.2025	Konzept vertiefen: Schnittstellen & Umsetzungsszenario (Kapitel 4.3–4.4)
KW 34	18.08.2025 – 24.08.2025	Analyse & Bewertung: Zielabgleich, Alternativen, Herausforderungen (Kapitel 5)
KW 35	25.08.2025 – 31.08.2025	Verfassen Ausblick (Kapitel 6), Vorbereitung Zusammenführung aller Teile
KW 36	01.09.2025 – 07.09.2025	Zusammenführung aller Kapitel, Formatierungen, Übergänge
KW 37	08.09.2025 – 14.09.2025	Korrekturschleife, Einholen Feedback, wissenschaftl. Feinjustierung

KW 38	15.09.2025 – 21.09.2025	Finale Abgabevorbereitung, Eigenständigkeitserklärung, Upload
-------	-------------------------	---