

**Bereitstellung offener Daten aus öffentlichen Einrichtungen
in Deutschland durch Content-Management-Systeme
am Beispiel von Hochschuldaten und TYPO3**

Bachelor-Arbeit

zur Erlangung des Grades Bachelor of Science
im Online-Studiengang Medieninformatik
der VFH an der Technischen Hochschule Brandenburg

vorgelegt von

Marco Habich
Nienburg / Weser

Betreuer: Sebastian Kreideweiß, M.Sc.
Zweitgutachter: Prof. Dr. Vera Meister

Beginn: 18.04.2023
Abgabe: 11.07.2023

Nienburg / Weser, den 6. Juli 2023

Inhaltsverzeichnis

Abbildungsverzeichnis	III
Tabellenverzeichnis	IV
Codeverzeichnis	V
Medienverzeichnis	VI
Abkürzungsverzeichnis	VII
Glossar	VII
1 Einleitung	1
2 Grundlagen	2
2.1 Hochschuldaten	8
2.2 Ziel	8
3 Manuelle Recherche	9
3.1 Umfrage	14
3.2 Auswertung zur manuellen Recherche	22
4 Automatische Recherche	23
4.1 Code	25
4.2 LD & Files	30
4.3 LD & Property	33
4.4 Auswertung zur automatischen Recherche	37
5 Prototyp	39
5.1 CollegeOrUniversity	43
5.2 Eine Frage von TYPO3?	44
5.3 Verwendungsmöglichkeiten	45
5.4 Erweiterungsmöglichkeiten	49
6 Fazit	50
Literaturverzeichnis	52

Anhang

Code

Prototyp

Ehrenwörtliche Erklärung

Abbildungsverzeichnis

Abb. 1: 5-Star-Steps	3
Abb. 2: TYPO3 v12 Maintenance-Support-Timeline	4
Abb. 3: JSON-LD	5
Abb. 4: Turtle.....	5
Abb. 5: Turtle in JSON-LD	5
Abb. 6: RDFa	6
Abb. 7: RDFa in JSON-LD	6
Abb. 8: Microdata.....	6
Abb. 9: Microdata in JSON-LD.....	6
Abb. 10: Semantic Web	7
Abb. 11: Hochschuldaten	8
Abb. 12: Global - Resultat.....	9
Abb. 13: Global - Format.....	9
Abb. 14: THB - Vorstellung	10
Abb. 15: THB - Zahlen	10
Abb. 16: Uni Leipzig - Zahlen.....	10
Abb. 17: Uni Leipzig - Vorstellung.....	10
Abb. 18: THB - HTML.....	10
Abb. 19: Uni Leipzig - HTML.....	10
Abb. 20: Umfrage.....	14
Abb. 21: Verzeichnisbaum	25
Abb. 22: Ablauf.....	26
Abb. 23: LD & Files	30
Abb. 24: Filter.....	33
Abb. 25: LD & Property	33
Abb. 26: Hochschuldaten - JSON	39
Abb. 27: JSON Editor.....	45
Abb. 28: Static Template	45
Abb. 29: Employee & Student #1	47
Abb. 30: Employee & Student #2.....	47
Abb. 31: Faculty & Course #1	47
Abb. 32: Faculty & Course #2	47
Abb. 33: Map.....	48

Tabellenverzeichnis

Tab. 1: Historie - Auszug	7
Tab. 2: HTML & Syntaxen	11
Tab. 3: Extension - Suche	12
Tab. 4: Extension - WordPress	12
Tab. 5: Extension - Drupal.....	13
Tab. 6: Extension - TYPO3	13
Tab. 7: Umfrageergebnis - Hochschuldaten.....	16
Tab. 8: Umfrageergebnis - Domainübersicht	17
Tab. 9: Umfrageergebnis - Domain #1	17
Tab. 10: Umfrageergebnis - Domain #2	18
Tab. 11: Umfrageergebnis - Domain #3	18
Tab. 12: Umfrageergebnis - Domain #4	19
Tab. 13: Umfrageergebnis - Domain #5	19
Tab. 14: Umfrageergebnis - Ähnlichkeiten	20
Tab. 15: Umfrageergebnis - Merkmale.....	21
Tab. 16: Umgebung.....	23
Tab. 17: Sprache / Framework / Library.....	24
Tab. 18: Recherchearten.....	24
Tab. 19: Code - Bereiche	25
Tab. 20: LD & Files - Legende	31
Tab. 21: Duplikate - Ausreißer.....	32
Tab. 22: Duplikate - Vergleich	32
Tab. 23: LD & Property - Legende	34
Tab. 24: Generierungsvergleich - Syntax.....	35
Tab. 25: Generierungsvergleich - Beispiele	36
Tab. 26: Generierungsvergleich - Auswertung	37
Tab. 27: Generierungsvergleich - Course	37
Tab. 28: Generierungsvergleich - CollegeOrUniversity.....	38

Codeverzeichnis

Code 1: Blöcke.....	27
Code 2: Variablen	27
Code 3: Auslesen - Id+json	28
Code 4: Auslesen - otherSchemas	28
Code 5: Auslesen - validExtensions.....	28
Code 6: JSON - Variable.....	29
Code 7: JSON - Werte	29
Code 8: Script	29
Code 9: Prototyp	39
Code 10: Allgemeine Daten	40
Code 11: numberOfEmployees.....	40
Code 12: sameAs	40
Code 13: image.....	40
Code 14: address - geo - hasMap	41
Code 15: openingHoursSpecification.....	41
Code 16: hasOfferCatalog - faculty.....	41
Code 17: hasOfferCatalog - course	42
Code 18: Validator #1	43
Code 19: Validator #2	43
Code 20: Visualized	44
Code 21: Link in Map	49
Code 22: Prototyp - Erweiterung.....	49

Medienverzeichnis

- [Med_csv] Habich, Marco: URLs von TYPO3-Auftritten, 2023, in: code/uni_typo.csv (erstellt am 23.04.2023).
- [Med_ica] Habich, Marco: Ordner mit archivierten Dateien, 2023, in: code/website/icsArchiv (letzte Änderung am 22.06.2023).
- [Med_ics] Habich, Marco: Ordner mit Dateien inkludiert von index.php, 2023, in: code/website/ics (letzte Änderung am 28.06.2023).
- [Med_imp] Habich, Marco: Importdatei erstellt unis_table für [Med_ics], 2023, in: code/uni_import.py (letzte Änderung am 06.07.2023).
- [Med_ind] Habich, Marco: Startseite der Website, 2023, in: code/website/index.php (letzte Änderung am 28.06.2023).
- [Med_jso] Habich, Marco: JSON erzeugt durch [Med_imp], 2023, in: code/website/icsArchiv/unis_tableChart_until_20230622_165310.json (erstellt am 11.06.2023).
- [Med_man] Habich, Marco: URLs von TYPO3-Auftritten als Alternative zu [Med_csv], 2023, in: code/uni_typoManual.py (letzte Änderung am 23.06.2023).
- [Med_mid] Habich, Marco: Prototyp - mittlere Version, 2023, in: code/possible_LDmiddle.json (erstellt am 02.05.2023).
- [Med_pro] Habich, Marco: Prototyp - erweiterte Version, 2023, in: code/possible_LD.json (erstellt am 02.05.2023).
- [Med_res] Habich, Marco: Ordner mit Dateien - Resultate der Umfrage, 2023, in: code/umfrage/results (letzte Änderung am 15.06.2023).

Abkürzungsverzeichnis

BoDaoE	Bereitstellung offener Daten aus öffentlichen Einrichtungen
CMS	Content-Management-System
CSS	Cascading Style Sheets
CSV	Comma-separated Values
ECTS	European Credit Transfer and Accumulation System
HTML	Hypertext Markup Language
JS	JavaScript
JSON	JavaScript Object Notation
LD	Linked Data
Mbit	Megabit
PDF	Portable Document Format
PHP	Hypertext Preprocessor
RDF	Resource Description Framework
SS/WS	Sommersemester/Wintersemester
THB	Technische Hochschule Brandenburg
VFH	Virtuelle Fachhochschule
WLAN	Wireless Local Area Network
WWW	World Wide Web

Glossar

Browser	Programm, um eine Website zu betrachten
Code	Abfolge von Anweisungen in einer Programmiersprache
Extension	Komponente, um TYPO3 zu erweitern
localhost	Domainname von einem lokalen Rechner
macOS	Betriebssystem der Rechner (Desktop / Laptop) von Apple
Redakteur	Rolle/Nutzer, die/der Seiteninhalte bearbeiten kann
Slider	Bereich, der Informationen enthält und automatisch scrollt
TYPO3	Content-Management-System, frei verfügbar
Web	Synonym für World Wide Web
Website	Abrufbare Seite, lokal oder im Web

**„Tue es oder tue es nicht.
Es gibt kein Versuchen.“**

Star Wars: Episode V - Das Imperium schlägt zurück, 1980

1 Einleitung

Die Bereitstellung offener Daten aus öffentlichen Einrichtungen (BoDaoE) ist durch Gesetze [BMI_Od] formuliert und dient dem Zweck, dass Rohdaten in eindeutiger maschinenlesbarer Form bereitgestellt werden.

Im Rahmen dieser Arbeit soll identifiziert werden, wie und womit die Bereitstellung offener Daten optimiert werden kann. Dabei liegt das Hauptaugenmerk auf Hochschuldaten, die mittels TYPO3 verwaltet werden und der Öffentlichkeit im World Wide Web (WWW) zur Verfügung stehen.

Motivation

Die These ist, dass in einem Content-Management-System (CMS) ein erheblicher Datenschatz vorliegt, der jedoch nicht in eindeutiger und maschinenlesbarer Form für Drittanwendungen weiterverarbeitbar ist. Daten können somit nicht in einem Knowledge Graph für („Künstliche Intelligenz“-) Anwendungen integriert werden.

TYPO3 wird - aufgrund der Verbreitung [CMS_dow] - als CMS beispielhaft herangezogen. Ein Ziel ist, dass der Redakteur einer TYPO3-Instanz bei der Nutzung kaum spürbaren Mehraufwand hat, um offene Daten bereitzustellen.

Durch den Abruf von offenen Daten durch Dritte, z.B. in leicht zugänglichen Anwendungen, profitiert die Gesellschaft.

Methode

Der Autor investiert Zeit in eine Recherche und deren Auswertung, um einen Wissensstand aufzubauen. Im Anschluss werden Variationen aufgezeigt. Anhand eines Prototypen sollen die Erkenntnisse praxisnah implementiert werden und - wenn möglich - eine Visualisierung daraus hervorgehen.

Schreibstil

Zugunsten einer schnellen Lesbarkeit wird auf eine Genderschreibweise verzichtet. Es wird aber ausdrücklich darauf hingewiesen, wenn es nicht explizit anders aufgeführt ist, dass alle Menschen dieser Welt gemeint sind.

2 Grundlagen

Es werden Definitionen von Begriffen und Produkten genannt, die im Zusammenhang mit dieser Arbeit stehen. Zudem wird am Ende des Kapitels der Begriff Hochschuldaten - im Sinne dieser Thesis - umschrieben.

Offene Daten

Auch *Open Data* genannt, beschreibt die Bundesregierung in ihrer Open-Data-Strategie [BMI_Os] mit prägnanten Sätzen:

[...] bieten weitreichende Nutzungspotenziale und spielen damit im nationalen und internationalen Datenökosystem eine vielversprechende, eigenständige Rolle [...] Nach der Definition der Richtlinie der EU über offene Daten und die Weiterverwendung von Informationen des öffentlichen Sektors (Richtlinie (EU) 2019/1024) sind Daten dann „offen“, wenn sie „von allen zu jedem Zweck frei verwendet, weiterverwendet und weitergegeben werden können“ [...] Zum Wohle einer lebendigen Demokratie und einer modernen Gesellschaft wollen wir die Qualität und Quantität von Verwaltungsdaten sowie ihre Bereitstellung als Open Data signifikant erhöhen und über das nationale Metadatenportal GovData auffindbar machen. Die Open-Data-Strategie will als Initiative auch für eine verstärkte Bereitstellung von offenen Daten durch die Wirtschaft, Wissenschaft und die Zivilgesellschaft motivieren. Denn die Grundlage für Innovationen in Wirtschaft und Gesellschaft und für verbessertes, evidenzbasiertes Regierungshandeln sind offene Daten sowohl aus dem öffentlichen als auch dem privaten Bereich. Open Data ist daher ein prädestinierter Bereich für zukünftig verstärkten Austausch zwischen Staat und Unternehmen, Bürgerinnen und Bürger, der Zivilgesellschaft und der Forschung [...]

Open Government Data ist eine Kategorie von *Open Data*, wird aber im Rahmen dieser Arbeit als Synonym betrachtet. Im Leitfaden der Bertelsmann Stiftung [BSt_Lei] heißt es: „Offene Verwaltungsdaten [...] sind daher nicht personenbezogene Daten, die von öffentlichen Stellen gesammelt, erstellt oder bezahlt wurden und der Allgemeinheit kostenlos zur Verfügung gestellt werden.“

Offene Daten - Technischer Hintergrund

Allgemein betrachtet sind alle Informationen auf einer Website oder in einem PDF offene Daten. Aber technisch gesehen ist ein PDF die unterste Stufe des 5-Sterne-Modells, welches Sir Timothy John Berners-Lee definiert hat, der als Erfinder des *World Wide Web* bekannt ist.

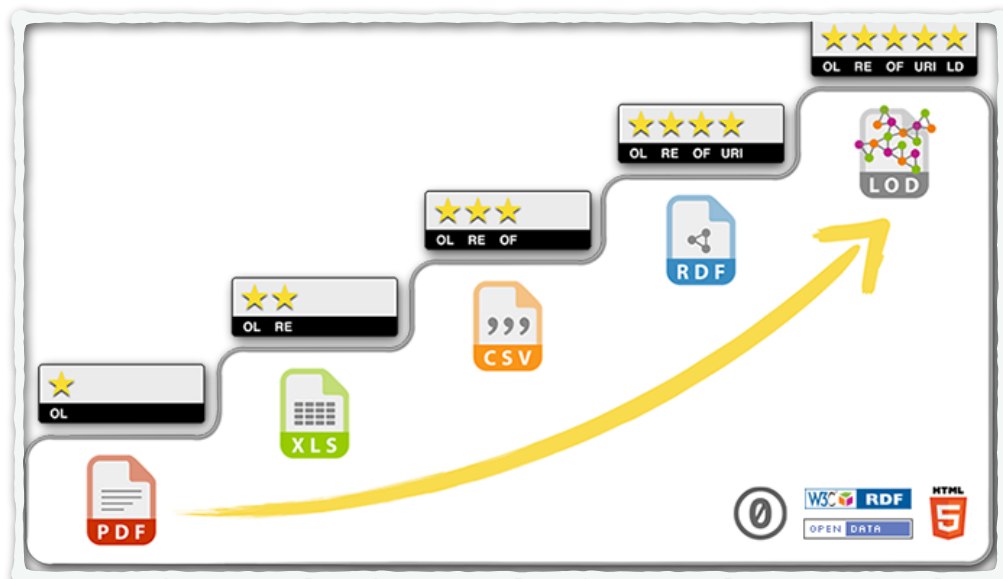


Abb. 1: 5-Star-Steps
Quelle: <https://5stardata.info/en>, 28.04.2023

In Abb. 1 ist ersichtlich, dass jede Folgestufe, die vorherigen Kriterien enthält. Die unterste Stufe bezeichnet solche Daten, die im Web unter einer „Open License“ (OL) bereitgestellt werden. Es folgt die Maschinenlesbarkeit (RE) und das offene nicht proprietäre Format (OF). Beinhalten die Daten ebenfalls Identifier (URI) und sind darüber hinaus mit anderen Daten verlinkt (LD) spricht man von 5-Sterne-Daten.

Content-Management-System

Mit Stand von Mitte 2022, als das CMS-Wahlpflichtmodul vom Autor belegt wurde, war laut *cmsmatrix.org* die Wahl unter eintausendfünfhundert Systemen möglich. Mindestens seit April 2023 hat die Website ihren Betrieb eingestellt. Weitere Recherchen ergaben unterschiedliche Zahlen. So liefert ein Marktreport [W3T_Rep] von *w3techs.com* eine Anzahl von 889 CMS. Dahingegen spricht der Hosting- und Cloud-Anbieter *ionos.de* in seinem CMS-Vergleich [ION_Ver] für das Jahr 2022 von zweihundertfünfzig bis dreihundert verschiedenen Programmen, die auf dem Markt verfügbar sind.

TYPO3

Die Geschichte reicht bis in das Jahr 1997 zurück, als Kasper Skårhøj die erste TYPO-Version entwickelte. Drei Jahre später erschien eine öffentliche Beta-Version und 2001 war die dritte Version von TYPO so erfolgreich, dass der Name, inklusive seiner Versionsnummer 3, zu einer Marke wurde.

TYPO3 basiert auf PHP, ist datenbankfähig und Open Source. Mit der Mehrsprachen- und Mandantenfähigkeit, sowie der Erweiterbarkeit mittels Extensions, sollen nur drei Faktoren genannt werden, die TYPO3 vorantreiben. Das CMS kann für jegliche Art von Internetauftritt genutzt werden. Auch wenn der weltweite Anteil bei 0,6% liegt [W3T_Usa], ist der Marktanteil im deutschsprachigen Raum besonders hoch [ION_Ver]. Der öffentliche Sektor spielt dabei ebenso eine größere Rolle. Im April 2023 ist mit TYPO3 v12 LTS die aktuellste Version auf dem Markt erschienen.

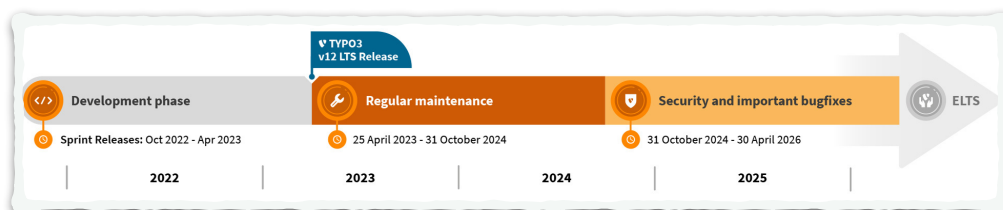


Abb. 2: TYPO3 v12 Maintenance-Support-Timeline
Quelle: <https://typo3.org/article/typo3-v12-lts-here-to-save-the-day>, 25.04.2023

JSON-LD

JSON (JavaScript Object Notation) wird im Web größtenteils zum Austausch von Daten genutzt. Viele Schnittstellen (API) können auf diese Art und Weise in einem strukturierten Format die gewünschten Daten liefern bzw. empfangen.

LD (Linked Data) sind strukturierte Daten, die wiederum Verbindungen zu anderen Daten in der gleichen Struktur oder einer Unterebene aufweisen. Außerdem gibt es die Möglichkeit, dass diese Datenverbindungen nicht im gleichen Kontext übermittelt werden müssen, sondern durch Identifier (URI) eine Integrität gewahrt bleibt.

W3C (World Wide Web Consortium) entwickelte dieses Format und bildete somit einen Container für verlinkte Daten. JSON-LD kann ohne viel Aufwand in HTML eingebettet werden. So können Suchmaschinen und nützliche Programme dieses Element auslesen.

JSON-LD (Beispiel)

Speichert man den Inhalt in einer Datei wird die Endung „jsonld“ bevorzugt. Beim MIME-Type (Multipurpose Internet Mail Extensions), im Sprachgebrauch auch „Content Type“ genannt, wird „application/ld+json“ verwendet.

```
{
  "@context": "https://json-ld.org/contexts/person.jsonld",
  "@id": "http://dbpedia.org/resource/John_Lennon",
  "name": "John Lennon",
  "born": "1940-10-09",
  "spouse": "http://dbpedia.org/resource/Cynthia_Lennon"
}
```

Abb. 3: JSON-LD
Quelle: <https://json-ld.org>, 15.05.2023

Anhand der vorherigen Abbildung erkennt man eine klare Struktur. In einer Enzyklopädie [Wiki_JL] wird die Art als konkrete Syntax von RDF beschrieben, siehe gleichnamigen Absatz auf der Folgeseite. Die Umschreibung „konkret“ bezieht sich in diesem Fall auf die Zerlegung und Darstellung der Daten.

Turtle

Terse RDF Triple Language [W3C_ttl] ermöglicht es, einen RDF-Graphen vollständig in einer kompakten und natürlichen Textform zu schreiben. Dabei ist die Kompatibilität zu N-Triples gegeben. Als Dateiendung nutzt man „ttl“ und der MIME-Type ist „text/turtle“.

Die jeweils abweichende Struktur - trotz gleicher Daten - kann man den Folgeabbildungen entnehmen.

EXAMPLE 154: A set of statements serialized in Turtle

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .

<http://manu.sporny.org/about#manu> a foaf:Person;
  foaf:name "Manu Sporny";
  foaf:homepage <http://manu.sporny.org/> .
```

Abb. 4: Turtle
Quelle: <https://www.w3.org/TR/json-ld/#turtle>, 17.05.2023

EXAMPLE 155: The same set of statements serialized in JSON-LD

```
{
  "@context": {
    "foaf": "http://xmlns.com/foaf/0.1/"
  },
  "@id": "http://manu.sporny.org/about#manu",
  "@type": "foaf:Person",
  "foaf:name": "Manu Sporny",
  "foaf:homepage": { "@id": "http://manu.sporny.org/" }
}
```

Abb. 5: Turtle in JSON-LD
Quelle: <https://www.w3.org/TR/json-ld/#turtle>, 17.05.2023

RDF / RDFa

Das Resource Description Framework (RDF) ist eine Sprache zur Darstellung von Informationen über Ressourcen im Web [W3C_rdf]. Dabei handelt es sich bei Ressourcen unter anderem um Dokumente, Personen und Objekte.

Die Ergänzung mittels „a“ bedeutet „Attribute“ [W3C_rdfa]. Somit besteht die Möglichkeit HTML zu erweitern und Daten einzubinden.

Für dieses Format sollen die Unterschiede nicht minder dargestellt sein.

```
EXAMPLE 162: RDFa fragment that describes three people
<div prefix="foaf: http://xmlns.com/foaf/0.1/">
  <ul>
    <li typeof="foaf:Person">
      <a property="foaf:homepage" href="http://example.com/bob/">
        <span property="foaf:name">Bob</span>
      </a>
    </li>
    <li typeof="foaf:Person">
      <a property="foaf:homepage" href="http://example.com/eve/">
        <span property="foaf:name">Eve</span>
      </a>
    </li>
    <li typeof="foaf:Person">
      <a property="foaf:homepage" href="http://example.com/manu/">
        <span property="foaf:name">Manu</span>
      </a>
    </li>
  </ul>
</div>
```

Abb. 6: RDFa

Quelle: <https://www.w3.org/TR/json-ld/#rdfa>, 17.05.2023

```
EXAMPLE 163: Same description in JSON-LD (context shared among node objects)
{
  "@context": {
    "foaf": "http://xmlns.com/foaf/0.1/",
    "foaf:homepage": {"@type": "@id"}
  },
  "@graph": [
    {
      "@type": "foaf:Person",
      "foaf:homepage": "http://example.com/bob/",
      "foaf:name": "Bob"
    }, {
      "@type": "foaf:Person",
      "foaf:homepage": "http://example.com/eve/",
      "foaf:name": "Eve"
    }, {
      "@type": "foaf:Person",
      "foaf:homepage": "http://example.com/manu/",
      "foaf:name": "Manu"
    }
  ]
}
```

Abb. 7: RDFa in JSON-LD

Quelle: <https://www.w3.org/TR/json-ld/#rdfa>, 17.05.2023

Microdata

Wird ebenfalls dazu verwendet, um Inhalt auf einer Website mit maschinenlesbaren Informationen anzureichern [WWG_md] und entstand mit HTML5.

Anhand von zwei Bildern sind die Abweichungen erkennbar.

```
-EXAMPLE 164: HTML that describes a book using microdata
<dl itemscope
  itemType="http://purl.org/vocab/frbr/core#Work"
  itemId="http://purl.org/works/45UB0JGZ5QKH8M">
  <dt-title/dt>
    <dd-cite itemprop="http://purl.org/dc/elements/1.1/title">Just a Geek</cite></dd>
  <dt-Byc/dt>
    <dd-span itemprop="http://purl.org/dc/elements/1.1/creator">Wil Wheaton</span></dd>
  <dt-Format/dt>
    <dd itemprop="http://purl.org/vocab/frbr/core#realization"
      itemscope
      itemType="http://purl.org/vocab/frbr/core#Expression"
      itemId="http://purl.org/products/9788596807683.EBOOK">
    <link itemprop="http://purl.org/dc/elements/1.1/type" href="http://purl.org/pro
      Print
    </dd>
    <dd itemprop="http://purl.org/vocab/frbr/core#realization"
      itemscope
      itemType="http://purl.org/vocab/frbr/core#Expression"
      itemId="http://purl.org/products/978859682189.EBOOK">
    <link itemprop="http://purl.org/dc/elements/1.1/type" href="http://purl.org/pro
      Ebook
    </dd>
  </dt>
```

Abb. 8: Microdata

Quelle: <https://www.w3.org/TR/json-ld/#microdata>, 17.05.2023

```
EXAMPLE 165: Same book description in JSON-LD (avoiding contexts)
{
  "id": "http://purl.org/works/45UB0JGZ5QKH8M",
  "type": "http://purl.org/vocab/frbr/core#Work",
  "http://purl.org/dc/elements/1.1/title": "Just a Geek",
  "http://purl.org/dc/elements/1.1/creator": "Wil Wheaton",
  "http://purl.org/vocab/frbr/core#realization":
  [
    {
      "id": "http://purl.org/products/9788596807683.EBOOK",
      "type": "http://purl.org/vocab/frbr/core#Expression"
    }, {
      "id": "http://purl.org/products/978859682189.EBOOK",
      "type": "http://purl.org/vocab/frbr/core#Expression"
    }
  ],
  "http://purl.org/dc/elements/1.1/type": {"@id": "http://purl.org/product-types"}
}
```

Abb. 9: Microdata in JSON-LD

Quelle: <https://www.w3.org/TR/json-ld/#microdata>, 17.05.2023

Historie - Auszug

MM / JJJJ	Syntax	Art / Link
10 / 2008	RDFa	W3C Recommendation
01 / 2014	JSON-LD 1.0	W3C Superseded Recommendation
02 / 2014	RDF 1.1 Turtle	W3C Recommendation
12 / 2014	Microdata to RDF	W3C Interest Group Note
03 / 2015	RDFa 1.1 Primer	W3C Working Group Note
07 / 2020	JSON-LD 1.1	W3C Recommendation
05 / 2023	RDF 1.2 Turtle	W3C Working Draft

Tab. 1: Historie - Auszug
Quelle: eigene Darstellung, 18.05.2023

Schema

Alle bereits genannten Syntaxen (JSON-LD, Turtle, RDFa, Microdata) benötigen vorzugsweise ein einheitliches Vokabular, damit Daten optimal ausgelesen werden können und somit eine Garantie besteht, dass die Kennzeichnung exakt dem Bereich zugeordnet ist, den man beschreiben möchte.

Aus dieser Idee und der Kooperation von Suchmaschinen (Google, Bing, Yahoo, Yandex) wurde 2011 durch die Initiative *Schema.org* eine erste Version veröffentlicht. Weiterentwicklungen, auch auf Grundlage von Diskussionen in öffentlichen Mailinglisten, sorgten für den Fortbestand. Release 21.0 erschien Ende Mai 2023 [Sch_Rel].

Semantic Web

Die Vereinigung der beschriebenen Elemente interpretiert der Student als *Semantic Web*. Durch eine einheitliche Wortwahl in den Formaten wird der einfache Austausch von Daten ermöglicht. Schlussfolgernd, durch die Maschinenlesbarkeit, ist eine sinnvolle Auswertung gegeben. Graphen können erzeugt werden, die durch ihre Knoten Verbindungen schaffen und Informationen bereitstellen. Es entsteht das „Web of data“ [W3C_Sem].

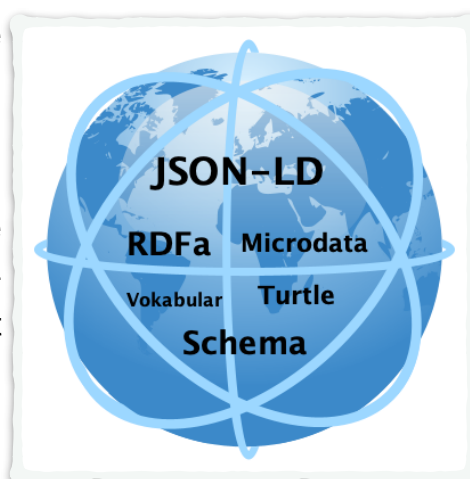


Abb. 10: Semantic Web
Quelle: eigene Darstellung, 17.05.2023

2.1 Hochschuldaten

Schaut man auf die Internetpräsenz einer Universität oder Hochschule findet man viele Informationen zum und um das Studium. Ähnliche Strukturen, beispielsweise die Auswahl von Fachbereichen und Studiengängen, sind immer vorhanden. Darüber hinaus gibt die Website Auskunft zu Kontaktmöglichkeiten, historischen Hintergründen und aktuellen Events.

Downloads - meist als PDF - erweitern das Webangebot. Hierbei handelt es sich zum Beispiel um hilfreiche Anmeldeformulare, aber auch Studieninhalte, die der Webpräsenz ähneln und kaum einen Mehrwert bieten.

Aufgrund der Datenvielfalt soll sich in dieser Thesis zunächst nur auf einen Bruchteil konzentriert werden.

Das Hauptaugenmerk liegt auf den Kontaktdaten, dem Gründungsdatum, sowie der Anzahl von Fachbereichen, Studiengängen, Professoren und Studenten.

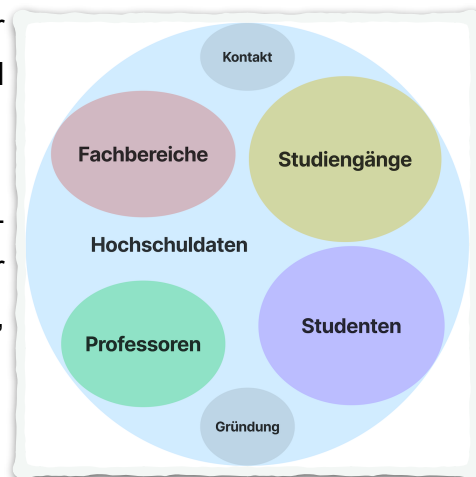


Abb. 11: Hochschuldaten
Quelle: eigene Darstellung, 18.05.2023

2.2 Ziel

Als Ergebnis der Arbeit soll die Zusammenführung der Hochschuldaten in einer geeigneten Syntax erfolgen. Resultierend daraus können Suchmaschinen oder Apps diese Daten wertvoller bereitstellen, wovon wiederum Nutzer profitieren.

3 Manuelle Recherche

Anhand von manuellen Recherchen soll sich der Aufgabenstellung genähert werden, um Ansätze für eine automatische Prüfung zu gestalten und die anschließende Realisierung zu optimieren. Bei allen Erhebungen wurden, wenn nicht gesondert gekennzeichnet, TYPO3-Internetseiten betrachtet.

Globale Verwendung

Das *WDC Project* [WDC_Pro] führt jährlich eine Statistik zum Thema „Structured Data“ durch. Die Zahlen aus dem Jahr 2022 können den beiden Abbildungen entnommen werden.

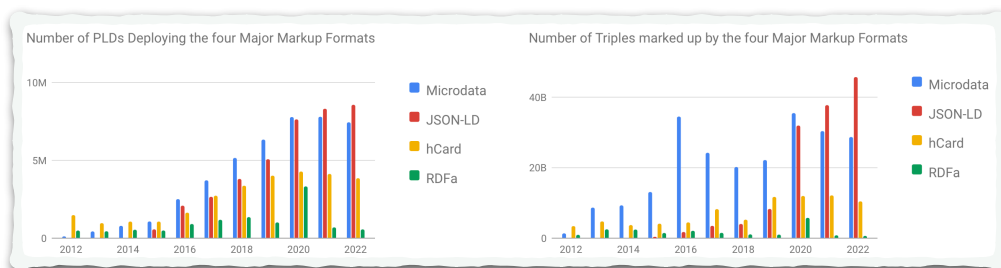


Abb. 12: Global - Resultat
 Quelle: <https://webdatacommons.org/structureddata/#results-2022-1>, 19.05.2023



Abb. 13: Global - Format
 Quelle: <https://webdatacommons.org/structureddata/2022-12/stats/stats.html>, 19.05.2023

Website

Mit der THB (Technische Hochschule Brandenburg) und der Universität Leipzig sind Beispiele genannt, die jeweils auf verschiedenen Seiten unterschiedliche Werte aufweisen.

2.600 Studierende, 67 Professorinnen und Professoren Hochschule Brandenburg ist eine junge und moderne Hochschule. Fachkräftesicherung und Technologietransfer machen von derzeit zehn Hochschulen für angewandte Wissenschaften. Wissenstransfer in die regionale Wirtschaft zu stärken. "Sicherheitsforschung" auf der Forschungslandkarte der

Abb. 14: THB - Vorstellung
Quelle: <https://www.th-brandenburg.de/hochschule/vorstellung-ueber-uns>, 18.05.2023

Zahlen / Daten / Fakten

Die Technische Hochschule Brandenburg auf einen Blick:

Gegründet 1992
2.467 Studierende
56 Professorinnen und Professoren
24 Studiengänge
3 Fachbereiche
60 internationale Partnerschaften

Stand: 01.02.2023
Die Daten werden jährlich aktualisiert.

Abb. 15: THB - Zahlen
Quelle: <https://www.th-brandenburg.de/hochschule/vorstellung-ueber-uns/zahlen-daten-fakten>, 18.05.2023



Abb. 17: Uni Leipzig - Vorstellung
Quelle: <https://www.uni-leipzig.de/#c56>, 18.05.2023

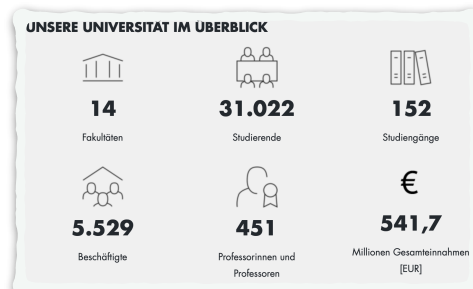


Abb. 16: Uni Leipzig - Zahlen
Quelle: <https://www.uni-leipzig.de/universitaet/profil/zahlen-und-fakten>, 18.05.2023

Die Zahlen (Abb. 15 & 16) besitzen folgende HTML-Passage (Abb. 18 & 19).

```
<div id="c140b9" class="csc-default frame-default frame-type-textpic frame-layout-0">
  <div id="h14059" class="zahlen / Daten / Fakten" />
  <div class="ce-textpic ce-right ce-intent">
    <div class="ce-gallery" data-ce-columns="1" data-ce-images="1"></div>
    <div class="ce-bodytext">
      <p>Die Technische Hochschule Brandenburg auf einen Blick:</p>
      <ul class="list">
        <li>Gegründet 1992</li>
        <li>2.467 Studierende</li>
        <li>56 Professorinnen und Professoren</li>
        <li>24 Studiengänge</li>
        <li>3 Fachbereiche</li>
        <li>60 internationale Partnerschaften</li>
      </ul>
      <p>Stand: 01.02.2023</p>
      <p>Die Daten werden jährlich aktualisiert.</p>
    </div>
  </div>
  </div>
  </div>
```

Abb. 18: THB - HTML
Quelle: <https://www.th-brandenburg.de/hochschule/vorstellung-ueber-uns/zahlen-daten-fakten>, 19.05.2023

```
<div class="module m128 gridelements_p11">
  <h2 class="h2">Unsere Universität im Überblick</h2>
  <div id="c11592" class="container-fluid">
    <div class="imageTextRow row smallSlick items-6">
      <div class="item col-md-4">
        
        <p>Fakultäten</p>
      </div>
      <div class="item col-md-4">
        
        <p>Studierende</p>
      </div>
      <div class="item col-md-4">
        
        <p>Studiengänge</p>
      </div>
      <div class="item col-md-4">
        
        <p>Beschäftigte</p>
      </div>
      <div class="item col-md-4">
        
        <p>Professorinnen und Professoren</p>
      </div>
      <div class="item col-md-4">
        <img alt="Euro symbol representing revenue" data-bbox="753 471 783 488"/>
        <p>Millionen Gesamteinnahmen [EUR]</p>
      </div>
    </div>
  </div>
```

Abb. 19: Uni Leipzig - HTML
Quelle: <https://www.uni-leipzig.de/universitaet/profil/zahlen-und-fakten>, 19.05.2023

HTML & Syntaxen

Stichprobenartige Kontrollen ergaben die Erkenntnis, dass bereits verschiedene Syntaxen, bezugnehmend auf Hochschuleseiten, genutzt werden.

CMS	Uni / Hochschule	Quelle	Syntax	Information
WordPress	CODE	Startseite	JSON-LD	<ul style="list-style-type: none"> - Website URL - Social Media URLs - Postanschrift - Mail, Telefon - Gründungsdaten
WordPress	Touro University	Startseite	JSON-LD	<ul style="list-style-type: none"> - Website URL - Facebook URL
Drupal	Augsburg	Startseite	<i>keine</i>	
Drupal	Bonn-Rhein-Sieg	Startseite	<i>keine</i>	
TYPO3	TH Brandenburg	Startseite	Microdata	<ul style="list-style-type: none"> - Article
TYPO3	TH Brandenburg	Zahlen / Daten / Fakten	Microdata	<ul style="list-style-type: none"> - BreadcumbList - Person
TYPO3	Universität Leipzig	Startseite	JSON-LD	<ul style="list-style-type: none"> - Website URL - Social Media URLs - Postanschrift - Telefon
TYPO3	FH Wedel	Startseite	JSON-LD	<ul style="list-style-type: none"> - Website URL
TYPO3	FH Wedel	Entdecken	JSON-LD	<ul style="list-style-type: none"> - BreadcumbList

Tab. 2: HTML & Syntaxen
Quelle: eigene Darstellung, 19.05.2023

CMS-Extension

Eine Extension - Bezeichnung in TYPO3 - ist eine Erweiterung, die das CMS optimiert oder spezielle Funktionen hinzufügt. Bei WordPress spricht man von Plugins und Drupal betitelt seine Erweiterungen als Module. Nutzt man die Suche (Link in Spalte „CMS“), mit den Schlagwörtern (Tabellenkopf in Tab. 3), erhält man unter der Vielzahl auch nützliche Ergebnisse.

CMS	Linked Data	Structured Data	Structured Content	Markup	Schema	JSON-LD
WordPress	6.618	1.724	2.947	1.584	749	194
Drupal	737	442	1.212	507	337	22
TYPO3	11	19	17	10	18	9

Tab. 3: Extension - Suche
Quelle: eigene Darstellung, 22.05.2023

Aus den drei genannten Systemen wurden jeweils drei themennahe Extensions ausgewählt und mit Stichpunkten umrissen.

CMS	Extension letzte Akt. Version	Bewertung Votings Installationen	Merkmale
WordPress	Yoast SEO 09.05.2023 20.7	4,8 27.574 5.000.000	- Meta-Tags, Sitemaps (Extensible Markup Language), Website-Breadcrumbs - tiefe <i>Schema.org</i> -Integration, die Suchmaschinen hilft, Inhalte besser zu verstehen
WordPress	Schema 13.08.2022 1.7.9.5	4,4 212 60.000	- generiert JSON-LD Output
WordPress	Schema App Structured Data 27.04.2023 1.22.0	3,9 35 10.000	- generiert JSON-LD Output

Tab. 4: Extension - WordPress
Quelle: eigene Darstellung, in Anlehnung an die Extension (Link in Spalte „Extension“), 22.05.2023

CMS	Extension letzte Akt. Version	Votings Nutzer / Woche	Merkmale
Drupal	Easy Breadcrumb 12.01.2023 2.0.5	106 44.299	- Metadata werden JSON-LD hinzugefügt
Drupal	RDF 21.10.2022 2.1.1	6 8.115	- stellt RDF auch in Drupal 10 zur Verfügung
Drupal	JSLD 19.07.2022 8.x-1.2	6 18	- hilft dabei JSON-LD Bereiche zu organisieren

Tab. 5: Extension - Drupal

Quelle: eigene Darstellung, in Anlehnung an die Extension (Link in Spalte „Extension“), 23.05.2023

CMS	Extension letzte Akt. Version	Votings Downloads	Merkmale
TYPO3	cs_seo 15.03.2023 8.0.0	15 183.039	- erweitert die Seiteneinstellung um Structured Data (JSON-LD)
TYPO3	schema 22.05.2023 2.8.0	13 127.308	- leichte Integration von Structured Data
TYPO3	virtuallocation 06.10.2022 1.3.0	0 3.285	- <i>schema</i> -Erweiterung, um das Element „Virtual Location“

Tab. 6: Extension - TYPO3

Quelle: eigene Darstellung, in Anlehnung an die Extension (Link in Spalte „Extension“), 23.05.2023

3.1 Umfrage

Im Zeitraum vom 20.05.2023 bis 06.06.2023 wurde eine eigens entwickelte Umfrage auf dem THB-Account veröffentlicht und den Kommilitonen angeboten.

Umfrage

☰

Was verbindest du mit dem Begriff Hochschuldaten?	Hochschuldaten
Deine Hochschul-Website lautet (Domain)?	Domain
... hat strukturierte Bereiche?	<input type="radio"/> ++ <input type="radio"/> + <input type="radio"/> % <input type="radio"/> - <input type="radio"/> --
... bietet dir , was du suchst?	<input type="radio"/> ++ <input type="radio"/> + <input type="radio"/> % <input type="radio"/> - <input type="radio"/> --
... deine Hinweise , falls du welche hast!	Hinweise
Wie bist du (damals) auf deine jetzige Hochschule aufmerksam geworden?	damals
... und was hat dich bewogen (dort) zu studieren?	Gründe
... hat die Hochschul-Website dazu beigetragen ?	<input type="radio"/> ++ <input type="radio"/> + <input type="radio"/> % <input type="radio"/> - <input type="radio"/> --
Kennst du andere gute Hochschul-Websites?	<input type="radio"/> ++ <input type="radio"/> + <input type="radio"/> % <input type="radio"/> - <input type="radio"/> --
... welche?	gute Websites
Kennst du andere schlechte Hochschul-Websites?	<input type="radio"/> ++ <input type="radio"/> + <input type="radio"/> % <input type="radio"/> - <input type="radio"/> --
... welche?	schlechte Websites
Wenn du andere Seiten kennst, gibt es Ähnlichkeiten bei der Bereichsstruktur?	<input type="radio"/> ++ <input type="radio"/> + <input type="radio"/> % <input type="radio"/> - <input type="radio"/> --
Planst du (zukünftig) andere Hochschulen anzusehen?	<input type="radio"/> ++ <input type="radio"/> + <input type="radio"/> % <input type="radio"/> - <input type="radio"/> --
Was ist dir bei einer Hochschul-Website wichtig ?	
... Infos zum Studium	<input type="radio"/> ++ <input type="radio"/> + <input type="radio"/> % <input type="radio"/> - <input type="radio"/> --
... Infos zum Studentenleben	<input type="radio"/> ++ <input type="radio"/> + <input type="radio"/> % <input type="radio"/> - <input type="radio"/> --
... Infos allgemeiner (Hochschul-) Natur	<input type="radio"/> ++ <input type="radio"/> + <input type="radio"/> % <input type="radio"/> - <input type="radio"/> --
... Infos zu Stadt und Umgebung	<input type="radio"/> ++ <input type="radio"/> + <input type="radio"/> % <input type="radio"/> - <input type="radio"/> --
... kurze Wege auf der Website	<input type="radio"/> ++ <input type="radio"/> + <input type="radio"/> % <input type="radio"/> - <input type="radio"/> --
... verfeinerte Suchmöglichkeiten	<input type="radio"/> ++ <input type="radio"/> + <input type="radio"/> % <input type="radio"/> - <input type="radio"/> --
... Bedienbarkeit / Handling	<input type="radio"/> ++ <input type="radio"/> + <input type="radio"/> % <input type="radio"/> - <input type="radio"/> --
... Schnelligkeit	<input type="radio"/> ++ <input type="radio"/> + <input type="radio"/> % <input type="radio"/> - <input type="radio"/> --
Platz für deine Bemerkungen!	Bemerkungen

SENDEN

Abb. 20: Umfrage
 Quelle: <https://informatik.th-brandenburg.de/~habich>, 20.05.2023

Umfrageergebnis

Damit die Studenten unbefangen an die Fragen gehen, hat der Autor den Titel der Thesis nicht genannt. Im genannten Zeitraum wurden zwei Plattformen angesprochen. Darunter - am 20.05.2023 - ein Forum auf *discord.com* mit rund zweihundert potenziellen Teilnehmern, sowie am 01.06.2023 das standort-übergreifende Forum aller VFH-Studenten auf: *moodle.oncampus.de*

Bei erstgenannter Wahl kamen lediglich vier Antworten innerhalb von zwei Tagen zustande. Der Start im Juni, mit rund dreißig Antworten in zwei Stunden, verlief deutlich besser. Zusammenfassend sind 78 Teilnahmen eingegangen, wovon 78 ausgewertet wurden.

Durch die selbstgeschriebene Umfrage, konnten die Ergebnisse [Med_res] zweckgebunden empfangen werden, um so auf verschiedene Auswertungsmöglichkeiten vorbereitet zu sein.

In Tab. 7 bis Tab. 15 ist - zur Übersichtlichkeit - die Website aus Abb. 20 aufgeteilt. Für alle Fragen, außer Textfelder, gilt:

++	positiv / zustimmend	+	Tendenz positiv / zustimmend
%	neutral		
-	Tendenz negativ / ablehnend	--	negativ / ablehnend

Da es keine Pflichtfelder in der Umfrage gab, bestand die Möglichkeit Fragen zu überspringen. Dieser Umstand ist mit „keine Angabe“ (kurz: kA) in der Auswertung gekennzeichnet.

Hinweis zu Tab. 7 Umfrageergebnis - Hochschuldaten

Aufteilung - der Reihenfolge nach - in drei Rubriken:

- themenbezogen
- uneindeutig
- am Thema vorbei (im Sinne der Definition laut Kapitel 2.1)

Textangaben wurden - teilweise - inhaltsgetreu gekürzt.

Auswertung in Prozent

Zahlen - gerundet auf zwei Nachkommastellen - in den Tabellen liefern den prozentualen Anteil, bezogen auf die Anzahl der Teilnahmen für die Rubrik.

Umfrageergebnis - Hochschuldaten

Teilnahmen - Gesamt	78
Teilnahmen - für diese Rubrik, zur prozentualen Anteil-Berechnung	78

Was verbindest du mit dem Begriff Hochschuldaten ?	Anteil
Allgemeine Informationen der Hochschule über sich selbst: <ul style="list-style-type: none"> - Kontaktdaten, Name, Standort, Größe, Statistische Kennzahlen - Studiengänge, Vorlesungsverzeichnis, Prüfungen, Struktur - Dozenten, Studenten, Organisatorisches, ASTA, Fachschaft 	19,23
So werden die Daten zum Hochschulbetrieb (Studierende, Fächer, Abschlüsse) genannt, die die Hochschulen an die für Bildung verantwortlichen Ministerien übermitteln, unter anderem auch: <ul style="list-style-type: none"> - Studierendenzahlen, Abbruchquoten, Studienzeiten - Zulassungszahlen, Erfolgsquote, Anmeldungen von Patenten 	5,13
<ul style="list-style-type: none"> - Studentennummern, Daten über Kommilitonen und Dozenten - Portale, Daten zur Hochschule wie Adresse - Dokumentation, Verwaltung, Skripte 	8,97
Daten die von der Hochschule gesammelt und verarbeitet werden	7,69
Information zur jeweiligen Hochschule, insbesondere Ort, Studiengänge, Zulassung, später dann Verwaltung	5,13
Alles rund um das Studium und die Hochschule. Sowohl für Studierende, als auch für Lehrende oder andere Mitarbeitende.	2,56
Personenbezogene Daten: <ul style="list-style-type: none"> - Name, Anschrift, Kontakt, Matrikelnummer, Studiengang, Noten - Anschreiben / Dokumentationen, Login, Moodle - alle für mein Studium und meine Leistungen notwendigen Daten - Daten, die sich im Laufe des Studiums auf dem Server ansammeln - Personalnummern von Dozenten - Abschlussarbeiten 	28,21
keine Angabe	15,38
Alle Daten, die von einer Hochschule selbst gespeichert werden: <ul style="list-style-type: none"> - Server-Inhalte, Aktenschränke - nur für die Hochschule gedachte Daten - Metadaten zu Prüfungen (für Prozessbeobachtung) 	7,69

Tab. 7: Umfrageergebnis - Hochschuldaten
Quelle: eigene Darstellung, 06.06.2023

Umfrageergebnis - Domainübersicht

Teilnahmen - Gesamt	78
Teilnahmen - für diese Rubrik	78

#	Domain	Anzahl	Anteil
1	www.th-brandenburg.de	16	20,51
2	www.hs-emden-leer.de	11	14,10
3	www.ostfalia.de	11	14,10
4	www.bht-berlin.de	10	12,82
5	www.th-luebeck.de	8	10,26
6	keine Angabe	7	8,97
7	www.fh-kiel.de	6	7,69
8	www.jade-hs.de	5	6,41
9	www.ash-berlin.eu	3	3,85
10	www.frankfurt-university.de	1	1,28

Tab. 8: Umfrageergebnis - Domainübersicht
Quelle: eigene Darstellung, 15.06.2023

In Tab. 9 - Tab. 13 werden die Hochschulen separat ausgewertet, die in Tab. 8 mindestens einen Anteil von 10% vorweisen können.

www.th-brandenburg.de						Anzahl
Teilnahmen - für Domain, zur prozentualen Options-Berechnung						16
	++	+	%	-	--	kA
... hat strukturierte Bereiche?	31,25	62,50		6,25		
... bietet dir , was du suchst?	25,00	56,25	6,25	6,25		6,25
... deine Hinweise , falls du welche hast!	<ul style="list-style-type: none"> - könnte optisch moderner sein - teilweise gleiche Daten auf verschiedenen Seiten - Suche nur bedingt hilfreich 					81,25
Wie bist du (damals) auf deine jetzige Hochschule aufmerksam geworden?	<ul style="list-style-type: none"> - Onlinesuche, Radio, Zeitung - Kollegen, Freunde 					6,25
... und was hat dich bewogen (dort) zu studieren?	<ul style="list-style-type: none"> - Flexibilität, aufgrund „Online“ - günstige Mieten, Nähe Berlin - viel Natur 					6,25
... hat die Hochschul-Website dazu beigetragen ?	31,25	18,75	31,25	6,25	6,25	6,25

Tab. 9: Umfrageergebnis - Domain #1
Quelle: eigene Darstellung, 15.06.2023

www.hs-empden-leer.de						Anzahl
Teilnahmen - für Domain, zur prozentualen Options-Berechnung						11
	++	+	%	-	--	kA
... hat strukturierte Bereiche?	36,36	36,36	9,09	9,09		9,09
... bietet dir , was du suchst?	27,27	27,27	27,27	18,18		
... deine Hinweise , falls du welche hast!	<ul style="list-style-type: none"> - unstrukturiert - lange Ladezeiten - Suche langwierig 					63,64
Wie bist du (damals) auf deine jetzige Hochschule aufmerksam geworden?	<ul style="list-style-type: none"> - Onlinesuche - Kollegen, Freunde, Verwandte 					
... und was hat dich bewogen (dort) zu studieren?	<ul style="list-style-type: none"> - Flexibilität, aufgrund „Online“ - Nähe Wohnort - Jobwechsel, Quereinstieg 					
... hat die Hochschul-Website dazu beigetragen ?	18,18	36,36	9,09	36,36		

Tab. 10: Umfrageergebnis - Domain #2
Quelle: eigene Darstellung, 15.06.2023

www.ostfalia.de						Anzahl
Teilnahmen - für Domain, zur prozentualen Options-Berechnung						11
	++	+	%	-	--	kA
... hat strukturierte Bereiche?	36,36	36,36	18,18	9,09		
... bietet dir , was du suchst?		72,73	9,09	18,18		
... deine Hinweise , falls du welche hast!	<ul style="list-style-type: none"> - Struktur für jede Fakultät anders und deren Inhalten sind gekoppelt - Anträge nicht leicht zu finden 					81,82
Wie bist du (damals) auf deine jetzige Hochschule aufmerksam geworden?	<ul style="list-style-type: none"> - Onlinesuche - Standort wie Wohnort - Kollegen, Freunde, Verwandte 					9,09
... und was hat dich bewogen (dort) zu studieren?	<ul style="list-style-type: none"> - Flexibilität, aufgrund „Online“ - Programmangebot - Nähe Wohnort 					9,09
... hat die Hochschul-Website dazu beigetragen ?	9,09	18,18	18,18	27,27	27,27	

Tab. 11: Umfrageergebnis - Domain #3
Quelle: eigene Darstellung, 15.06.2023

www.bht-berlin.de						Anzahl
Teilnahmen - für Domain, zur prozentualen Options-Berechnung						10
	++	+	%	-	--	kA
... hat strukturierte Bereiche?	40,00	50,00	10,00			
... bietet dir , was du suchst?	10,00	50,00	30,00			10,00
... deine Hinweise , falls du welche hast!	<ul style="list-style-type: none"> - unübersichtlich - man muss viel klicken - eher Google als Website 					80,00
Wie bist du (damals) auf deine jetzige Hochschule aufmerksam geworden?	<ul style="list-style-type: none"> - Onlinesuche - Standort in der Region - Freunde, Empfehlung 					
... und was hat dich bewogen (dort) zu studieren?	<ul style="list-style-type: none"> - Flexibilität, aufgrund „Online“ - Weiterentwicklung - Nähe Wohnort 					
... hat die Hochschul-Website dazu beigetragen ?	10,00	20,00	10,00	20,00	40,00	

Tab. 12: Umfrageergebnis - Domain #4
Quelle: eigene Darstellung, 15.06.2023

www.th-luebeck.de						Anzahl
Teilnahmen - für Domain, zur prozentualen Options-Berechnung						8
	++	+	%	-	--	kA
... hat strukturierte Bereiche?	37,50	37,50	25,00			
... bietet dir , was du suchst?	50,00	37,50	12,50			
... deine Hinweise , falls du welche hast!	<ul style="list-style-type: none"> - Struktur wirkt etwas aus der Zeit gefallen - Hochschulseite ist strukturiert, Kurse/Infos sind gut zu finden 					62,50
Wie bist du (damals) auf deine jetzige Hochschule aufmerksam geworden?	<ul style="list-style-type: none"> - Onlinesuche - Standort in der Region - Empfehlung 					
... und was hat dich bewogen (dort) zu studieren?	<ul style="list-style-type: none"> - Flexibilität, aufgrund „Online“ - Nähe Wohnort - Interesse 					
... hat die Hochschul-Website dazu beigetragen ?	12,50	25,00	12,50	25,00	25,00	

Tab. 13: Umfrageergebnis - Domain #5
Quelle: eigene Darstellung, 15.06.2023

Umfrageergebnis - Ähnlichkeiten

Teilnahmen - Gesamt	78
Teilnahmen - für diese Rubrik	78

	++	+	%	-	--	kA
Kennst du andere gute Hochschul-Websites?	6,41	11,54	33,33	15,38	26,92	6,41
... welche?	ehb Berlin, HS Stralsund, IU Berlin, Online-Kurse der Uni Potsdam, TU Braunschweig, Uni Hamburg, Wismar, Wolfenbüttel, htw-berlin.de, jade-hs.de, ovgu.de, th-brandenburg.de, th-luebeck.de, uni-frankfurt.de, uni-halle.de, uni-leipzig.de, uni-trier.de					
Kennst du andere schlechte Hochschul-Websites?	8,97	5,13	30,77	10,26	32,05	12,82
... welche?	CAU Kiel, FH Westküste, HS Bremerhaven, HS Emden-Leer, HS Harz, Karlsruhe, TU Dresden, TU Braunschweig, leuphana.de, rub.de, th-luebeck.de, tu-ilmenau.de, uni-osnabrueck.de					
Wenn du andere Seiten kennst, gibt es Ähnlichkeiten bei der Bereichsstruktur?	5,13	19,23	26,92	5,13	7,69	35,90
Planst du (zukünftig) andere Hochschulen anzusehen?	7,69	10,26	8,97	21,79	42,31	8,97

Tab. 14: Umfrageergebnis - Ähnlichkeiten
Quelle: eigene Darstellung, 09.06.2023

Bei der letzten Frage wäre es interessant, ob der Antwortgeber einen Bachelor-/Masterstudiengang besucht oder andere Gründe vorliegen. Denn die „negative“ Stimmabgabe von 64,1 Prozent (50 von 78 Stimmen) ist augenfällig.

Umfrageergebnis - Merkmale

Teilnahmen - Gesamt	78
Teilnahmen - für diese Rubrik	78

Was ist dir bei einer Hochschul-Website wichtig ?	++	+	%	-	--	kA
... Infos zum Studium	92,31	5,13	1,28			1,28
... Infos zum Studentenleben	5,13	28,21	38,46	17,95	8,97	1,28
... Infos allgemeiner (Hochschul-) Natur	16,67	35,90	29,49	12,82	1,28	3,85
... Infos zu Stadt und Umgebung	3,85	23,08	32,05	17,95	20,51	2,56
... kurze Wege auf der Website	58,97	33,33	5,13			2,56
... verfeinerte Suchmöglichkeiten	66,67	23,08	6,41		1,28	2,56
... Bedienbarkeit / Handling	75,64	20,51				3,85
... Schnelligkeit	51,28	32,05	11,54	2,56		2,56
Platz für deine Bemerkungen!	<ul style="list-style-type: none"> - verbessertes WholsWho - klare Gliederung, wenige Ebenen, prägnante Infos - ausführliche Beschreibungen zum Ablauf des Studiums 					87,18

Tab. 15: Umfrageergebnis - Merkmale
Quelle: eigene Darstellung, 09.06.2023

Umfrageergebnis - Zusammenfassung

Dem Autor dieser Arbeit war bei der Umfrage insbesondere die Meinung anderer Studenten - zur Definition von Hochschuldaten - wichtig. So ergibt sich daraus, dass 24,36% den Begriff im Sinne dieser Thesis deuten. Eine nicht eindeutige Zuordnung lag bei 24,35%. Über ein Drittel (35,9%) meinten, dass es personenbezogene Daten oder interne Hochschulangelegenheiten sind. 15,38% der Teilnehmenden machten keine Angabe zu diesem Begriff.

Gute und schlechte Hochschul-Websites, sowie Einzelheiten zu bestimmten Domains, sind persönliche Meinungen. Jedoch kann man aus diesen Bereichen weitere Studien (Graphical User Interface, User Experience) in Gang setzen.

Stichwörter (Flexibilität, Jobwechsel, Quereinstieg, wohnortnah) waren präsent und sollten bei der späteren Umsetzung eine kleine Rolle einnehmen.

3.2 Auswertung zur manuellen Recherche

Bei der manuellen Recherche war erkennbar, dass JSON-LD die letzten vier Jahre einen großen Sprung gemacht hat, wie die hier aufgeführten Statistiken belegen. Die bis dato führenden Formate wurden deutlich überholt.

Betrachtet man die Nutzersicht, kann man zunächst vermuten, dass durch die vorhandenen Menüs auf den Seiten ein schneller Überblick möglich ist. Doch die Gefahr ist hoch, dass man sich bei der Anzahl der Folgeoptionen verirrt. Erschwerend kommt hinzu, dass jede Navigation unterschiedlich ist.

Slider, die eine Internetseite verschönern sollen, sind jedoch für den Bereich „Zahlen und Fakten“ deplatziert, wie anhand der FH Wedel auf der Startseite [FHW_Sli] unübersehbar ist. Zwar ist es möglich, diesen Slider durch eine Mausbewegung anzuhalten oder selbst zu steuern, doch eine Übersicht fehlt.

Vereinzelte Sichtungen, direkt im HTML-Bereich der Hochschulseite, ließen die Verwendung von strukturierten Daten erkennen. Dabei handelt es sich vermehrt um „BreadcrumbList“, also die Navigationsführung zur Seite, auf der man sich momentan befindet.

Auswertungen zur Umfrage - sowie eine Zusammenfassung - wurden auf den vorhergehenden Seiten aufgeführt. Ein wesentlicher Punkt war, dass nur knapp ein Viertel der Teilnehmenden den Begriff „Hochschuldaten“ in diesem Kontext richtig eingeschätzt hat.

Der Websitebesucher muss den Überblick wahren, da alle Informationen zu Fachbereichen oder Studiengängen auf verschiedenen Seiten sind. Mit hoher Wahrscheinlichkeit erhält er das gesuchte Material, doch die Zeitintervalle bis zur Ergebnissammlung hängen vom Know-how und dem Seitenumgang ab.

Eine Suchmaschine oder andere Anwendungen, welche die Präsenz durchsuchen, können nur durch einen intelligenten Algorithmus logische Ergebnisse liefern. Strukturierte Daten, vorzugsweise in Form von JSON-LD, bündeln Informationen und würden ein immenses Plus schaffen.

4 Automatische Recherche

Zugrunde liegen 225 TYPO3-Auftritte. Die dazugehörige Domain wurde über CMSsensus [CMS_dow] ermittelt. Auf dieser Plattform sind insgesamt 415 Einträge (Universities) gelistet. Die relevanten 225 URLs wurden in eine CSV-Datei [Med_csv] geschrieben, um sie in Automatisierungen einzubinden.

Im späteren Verlauf der Arbeit wurden diese URLs in eine Python-Variable gesetzt und mit weiteren Daten angereichert [Med_man].

Der Code für die automatische Recherche wurde vom Studenten geschrieben. Aufgrund seiner Erfahrung konnte er verschiedene Programmier-/Skript-sprachen kombinieren und visualisierte die Ergebnisse auf einer Website.

Umgebung

Im Anhang ist der Ordner „code“ mit dem darin enthaltenen Ordner „website“. Weist man Letztgenannten einem lokalen Webserver zu, startet diesen und öffnet die localhost-Adresse in einem Browser, sind die Bestandteile der Recherchen von diesem Kapitel und seiner Unterpunkte abrufbar. Tab. 16 listet die Komponenten auf, die für die Eigenentwicklung verwendet wurden.

	Art	Merkmale
System	Rechner	<ul style="list-style-type: none">- MacBook Pro (Retina 13“, 2012)- 2,5 GHz Dual-Core Intel Core i5- 8 GB 1600 MHz DDR3- macOS Catalina (Version 10.15.7)
Visual Studio Code	Entwicklungs- umgebung	<ul style="list-style-type: none">- Version 1.76.2 (Universal)- Datum 2023-03-14T17:54:09.061Z
MAMP	Webserver	<ul style="list-style-type: none">- Version 6.8 (1258)- Apache, PHP 8.2
Browser	Programm	<ul style="list-style-type: none">- Safari (Version 15.6.1)- Firefox (Version 113.0.2 / 64-Bit)- Chrome (Version 113.0.5672.126)- Opera (Version 74.0.3911.218)

Tab. 16: Umgebung
Quelle: eigene Darstellung, 27.05.2023

Sprache / Framework / Library

Sprache wird in der Folgetabelle als Synonym für Programmier-/Skriptsprache verwendet. Frameworks und Libraries, wie *Bootstrap* und *amCharts* sind Helfer, um erweiterte Funktionen auf einer Website bereitzustellen.

Sprache / Framework / Library	Merkmale / Beispielpfad in „code“
Python 3.9.7	<ul style="list-style-type: none"> - durchsucht Domains, erstellt PHP-Dateien - code/uni_import.py
PHP 8.2	<ul style="list-style-type: none"> - Abfragen von Bedingungen - in HTML-Kombination zur Visualisierung - code/website/index.php - code/website/ics/unis_tableChart.php
HTML5, CSS3, JS	<ul style="list-style-type: none"> - Visualisierung - code/website/index.php - code/website/css/style.css
JSON	<ul style="list-style-type: none"> - zusätzliche Speicherung von Inhalten für Kontrollen oder spätere Verwendungen - code/website/ics/unis_tableChart.json
Bootstrap 5.2.0	<ul style="list-style-type: none"> - CSS / JS - code/website/css - code/website/js
amCharts v5	<ul style="list-style-type: none"> - Erstellung von Charts - teilweise eigene Editierungen - code/website/js/amCharts5/indexChart.js

Tab. 17: Sprache / Framework / Library
Quelle: eigene Darstellung, 27.05.2023

Recherchearten

Der Code [Med_imp] ist so ausgelegt, dass der Nutzer mehrere Optionen hat. Zeitangaben beruhen auf den - in Tab. 16 - erwähnten Rechnerangaben, einer WLAN-Verbindung (300 Mbit, hinter einer 1000 Mbit-Internetverbindung), sowie einem Timeout bei der URL-Anfrage von zwei Sekunden.

Rechercheart	Option	Domains	~ Dauer
LD & Files		225	7h 45m
LD & Property		225	5h 15m
LD & Property	devModeOnlyIndex = 1	225	3m

Tab. 18: Recherchearten
Quelle: eigene Darstellung, 04.06.2023

4.1 Code

Ist der „website“-Ordner als Stammverzeichnis in einem Webserver zugewiesen, wird, nach Aufruf der localhost-Adresse in einem Browser, automatisch *index.php* [Med_ind] geöffnet. Diese inkludiert relevante Dateien aus dem ics-Ordner [Med_ics]. Zudem werden CSS- und JS-Elemente geladen.

Der Ordner „umfrage“ enthält die Website für die Umfrage inklusive Resultate, wie in Kapitel 3.1 erwähnt und wird deshalb hier nicht weiter betrachtet.

Bereich	Nutzen
css	Beinhaltet Stylesheets für Seiten
fonts	oder deren Bestandteile.
ics / content...	Texte für die Website.
ics / funcs_body.php	Wiederverwendbare Funktionen.
ics / head.php	Metadaten für Website.
ics / modals.php	Visualisierung der Popup-Fenster.
ics / tab...	Visualisierung der Website.
ics / unis_table...	Generiert durch: <i>uni_import.py</i>
icsArchiv	Führt man <i>uni_import.py</i> aus und hat in selbiger Datei die Variable „withArchiv“ auf eins gesetzt, werden Kopien der <i>unis_table</i> -Dateien (Vorgänger) angelegt.
img	Icons zur visuellen Unterstützung.
js	Ablageort von <i>Bootstrap</i> und <i>amCharts</i> , wobei für Letzteres weitere Dateien angelegt wurden.
index.php	Grundgerüst für Website.
uni_import.py	Erzeugt <i>unis_table</i> -Dateien.
uni_typo.csv	Deklaration von 225 Domains.
uni_typoManual.py	Als Alternative zur CSV-Datei mit weiteren Daten angereichert.

Tab. 19: Code - Bereiche
Quelle: eigene Darstellung, 10.06.2023

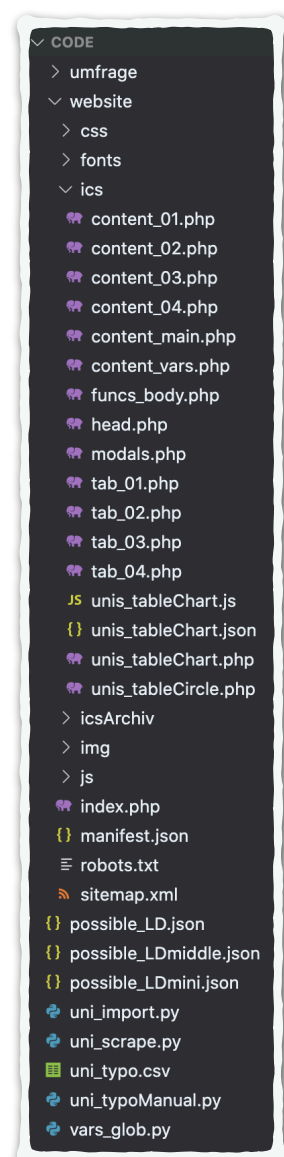


Abb. 21: Verzeichnisbaum
Quelle: Visual Studio Code, 10.06.2023

Ablauf

Ein Nutzer (engl.: User) öffnet die Website [Med_ind] und erhält verschiedene Ansichten, welche die Hochschuldaten repräsentieren. Dazu zählen Tabellen, Diagramme und eine Landkarte mit den Standorten. Darüber hinaus kann der Entwickler (engl.: Developer) eine neue Version der automatisch generierten Seiten erzeugen und somit die Hochschuldaten aktualisieren.

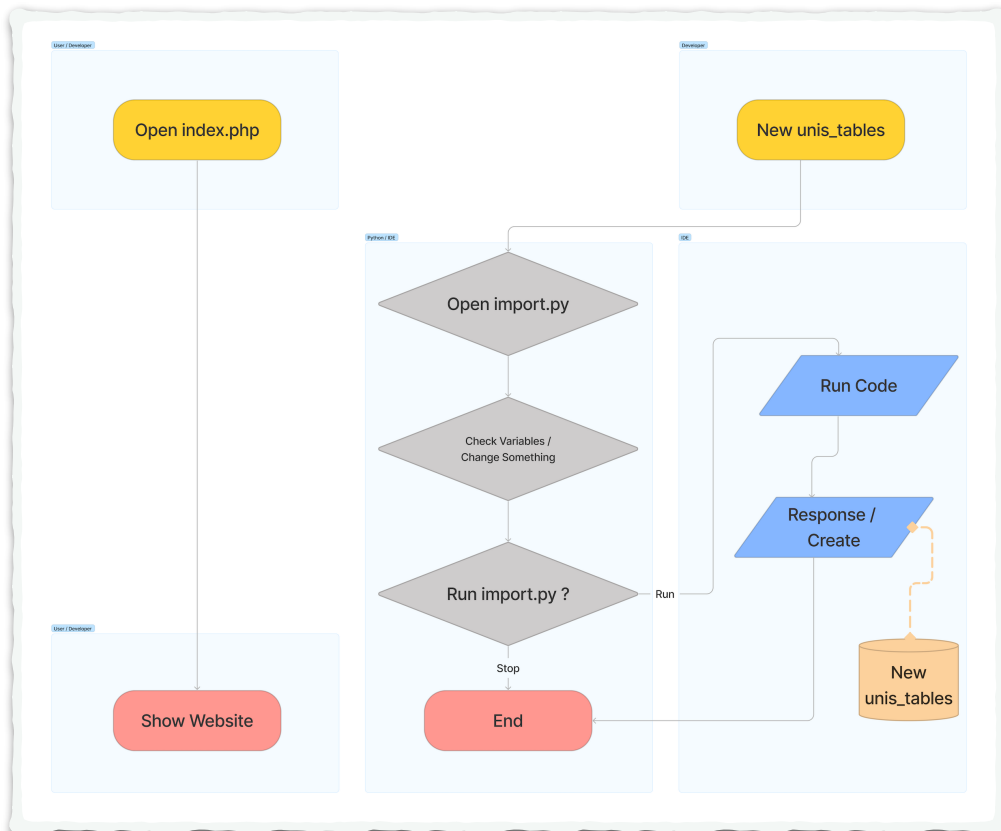


Abb. 22: Ablauf

Quelle: eigene Darstellung, erstellt mit www.figma.com, 14.06.2023

Mit „unis_tables“ sind in diesem Kontext PHP-Dateien gemeint. Es ist ebenso möglich die Daten in einer Datenbank abzulegen. Wobei in dem Fall die *index.php* angepasst werden muss, um wiederum die Datenbank auszulesen, damit der Nutzer die Daten angezeigt bekommt. Aus Performancegründen scheint die gewählte PHP-Lösung optimal zu sein.

Der Code ist flexibel gestaltet. Als Beispiel sei die Variable „withJson“ genannt. Sie erzeugt eine JSON-Datei mit Hochschuldaten. Demzufolge kann man die *index.php* dahingehend anpassen, dass das JSON ausgelesen wird. Dieses Mittel wurde für den Reiter „Versionen“ - auf der lokalen Website - angewendet.

uni_import.py

Die Datei lässt sich in Blöcke aufteilen und ist im Code gekennzeichnet.

- Deklaration von Variablen
- Import Universitäten und Hochschulen
- Lauf durch Universitäten und Hochschulen
- Schreiben der optionalen JS- und JSON-Datei
- Schreiben der PHP-Dateien (Tabellen, Charts, Circle, Map)

Im Gegensatz zu sonstigen Philosophien ist der Code hauptsächlich in dieser Datei verankert und kaum modular ausgegliedert.

```
# *****  
# generals (editable)  
# *****  
# generals (also editable, but take care what you do)  
# *****  
# declare uni_version  
# *****  
# (de)activate some stuff  
# *****  
# (de)activate some generals  
# *****  
# helpers  
# *****  
# import all unis  
# *****  
# scrape all unis  
# *****  
# write .js (ld+json)  
# *****  
# write .json  
# *****  
# write .php (chart)  
# *****  
# write .php (circle)  
# *****
```

Code 1: Blöcke
Quelle: [Med_imp], Visual Studio Code,
16.06.2023

Bestimmte Variablen können mit wenigen Handgriffen verändert werden.

```
withArchiv = 1 # 0 or 1 # if 1 save previous version  
randomMode = 0 # 0 or 1 # if 1 then random values of uniSource  
randomModeRecs = 3 # records for randomMode  
devMode = 0 # 0 or 1 # if 1 then special/shorter run  
devModeOnlyIndex = 0 # 0 or 1 # if 1 then special options  
devModeReduceTo = 6 # reduce uniSource
```

Code 2: Variablen
Quelle: [Med_imp], Visual Studio Code, 16.06.2023

Mittels aktiviertem „randomMode“ werden zufällige Universitäten und Hochschulen aus dem Import gewählt. Dies war bei Testläufen und für Recherchen hilfreich. Ferner wurde so eventuell ein IP-Adressen-Eintrag auf einer Blacklist vermieden, falls man eine Domain zu oft angesprochen hätte.

Insbesondere „devModeOnlyIndex“ hat, wie bei den Recherchearten erwähnt, enormen Einfluss auf die Zeitspanne des Laufes und wird deshalb in Kapitel 5.3 „Verwendungsmöglichkeiten“ wiederholt aufgegriffen.

Ein wichtiges Code-Element ist das Auslesen. Nach erfolgreichem Ansprechen der Website wird diese nach einem Muster (Regex) durchsucht und die Ergebnisse den Variablen zugewiesen.

```
# *****  
# ld+json  
result_j.clear()  
result_j = uni_scrape.reFindallWith(str(webText), '<script ', '</script>', '<script type=\"application/ld.+json.*?>.*?</script>')  
if len(result_j) > 0:  
    if result_j[0] == 'match':  
        scriptText = str(result_j[1])  
        if (result_script.count(scriptText) == 0):  
            result_script.append(scriptText)  
            result_scriptAll.extend(result_j[1])
```

Code 3: Auslesen - ld+json
Quelle: [Med_imp], Visual Studio Code, 16.06.2023

```
# *****  
# otherSchemas  
if withOtherSchemas == 1:  
    for o in otherSchemas:  
        result_j.clear()  
        if o == 'Microdata':  
            result_j = uni_scrape.reFindallWith(str(webText), '< ', '>', 'itemscope itemtype=.*?>')  
        elif o == 'RDFa':  
            result_j = uni_scrape.reFindallWith(str(webText), '< ', '>', 'vocab=.*? typeof=.*?>')  
        else:  
            pass  
        if len(result_j) > 0:  
            if result_j[0] == 'match':  
                scriptText = str(result_j[1])  
                if (result_scriptOtherSchemas.count(scriptText) == 0):  
                    result_scriptOtherSchemas.append(scriptText)  
                    result_scriptAll.append(str(web+e)) # extra line as marker # here append, all in one  
                    result_scriptAll.extend(result_j[1]) # here extend, because it's iterable  
                    oCount = result_otherSchemas.get(o, 0) + len(result_j[1])  
                    result_otherSchemas.update({o: oCount})
```

Code 4: Auslesen - otherSchemas
Quelle: [Med_imp], Visual Studio Code, 16.06.2023

```
# *****  
# validExtensions  
if withValidExtensions == 1:  
    try:  
        result_raw = BeautifulSoup(str(webText), 'html.parser')  
    except:  
        result_raw = BeautifulSoup("", 'html.parser')  
        # parseError_soup.append(web+e) # maybe for log, but only 1 link of over 36thousand  
        # type was []  
        # now: directly count and already here update json  
        peSo = result_extensions.get(keyPeSo, 0) + 1  
        result_extensions.update({keyPeSo: peSo})  
    for index, element in enumerate(result_raw.select('a')):  
        try:  
            tempLink = element.attrs['href'] # element.text  
            if ((tempLink.lower()).endswith(validExtensions)):  
                if (result_files.count(tempLink) == 0):  
                    result_files.append(tempLink)  
                    tempLinkExt = tempLink[tempLink.rfind("."):]  
                    for v in validExtensions:  
                        if tempLinkExt == v:  
                            vCount = result_extensions.get(v, 0) + 1  
                            result_extensions.update({v: vCount})  
                else:  
                    if devMode == 1 and len(uniSource) == 1:  
                        duplicatesArr.append(tempLink)  
                        duplicates = result_extensions.get(headDupl, 0) + 1  
                        result_extensions.update({headDupl: duplicates})  
        except KeyError:  
            pass
```

Code 5: Auslesen - validExtensions
Quelle: [Med_imp], Visual Studio Code, 16.06.2023

Die - durch das Auslesen - belegten Variablen finden unter anderem Anwendung in einem JSON-String, der optional in eine Datei geschrieben wird.

```
uni = {
  "name": uniName,
  "web": web,
  "webCode": webCode,
  "links": anzahlOrg,
  "linksReduce": anzahl,
  keyType: result_scriptTypes.copy(),
  keyExt: result_extensions.copy(),
  keyOs: result_otherSchemas.copy(),
  "circlesDataManual": circlesDataManual,
  "wantedData": [{
    "numberOfEmployees": [
      {
        "unitText": "professor",
        "value": professor
      },
      {
        "unitText": "staff",
        "value": staff
      },
      {
        "unitText": "student",
        "value": student
      }
    ],
    "hasOfferCatalog": [
      {
        "@type": "OfferCatalog",
        "name": "faculty",
        "numberOfItems": faculty,
        "itemListElement": []
      },
      {
        "@type": "OfferCatalog",
        "name": "course of study",
        "numberOfItems": courseOfStudy,
        "itemListElement": []
      }
    ]
  }],
  "geo": {
    geoCoords[0]: latitude,
    geoCoords[1]: longitude
  },
  "durationBegin": str(durationBegin),
  "durationEnd": str(durationEnd),
  "duration": str(duration)
}
```

Code 6: JSON - Variable
Quelle: [Med_imp], Visual Studio Code, 16.06.2023

```
{
  "name": "th-brandenburg",
  "web": "https://www.th-brandenburg.de",
  "webCode": 200,
  "links": 116,
  "linksReduce": 0,
  "ld": {
    "all": 0,
    "Course": 0,
    "Event": 0,
    "EducationEvent": 0,
    "LearningResource": 0,
    "OfferCatalog": 0,
    "WebSite": 0,
    "CollegeOrUniversity": 0,
    "EducationalOrganization": 0,
    "LocalBusiness": 0,
    "NGO": 0,
    "Organization": 0,
    "VirtualLocation": 0
  },
  "ext": {},
  "os": {
    "Microdata": 445,
    "RDFa": 0
  },
  "circlesDataManual": "found",
  "wantedData": [
    {
      "numberOfEmployees": [
        ...
      ],
      "hasOfferCatalog": [
        ...
      ]
    }
  ],
  "geo": {
    "latitude": "52.4118837",
    "longitude": "12.5367298"
  },
  "durationBegin": "2023-06-11 15:46:11.486387",
  "durationEnd": "2023-06-11 15:47:00.839419",
  "duration": "0:00:49.353032"
}
```

Code 7: JSON - Werte
Quelle: [Med_jso], Visual Studio Code, 16.06.2023

Für Charts, Circle und Map benötigt *amCharts* JS-Variablen. Dieser Umstand fand eine Lösung, indem innerhalb vom Python-Code der zu erzeugenden PHP-Datei ein script-Tag hinzugefügt wurde.

```
<script>
  var uniCircleEmployees = [{nameE: "Employees", childrenE: ['' + uniCircleEmployees + '']}]
  var uniCircleCatalog = [{nameC: "Catalog", childrenC: ['' + uniCircleCatalog + '']}]
  var uniGeo = ['' + uniGeo + '']
</script>
```

Code 8: Script
Quelle: [Med_imp], Visual Studio Code, 16.06.2023

4.2 LD & Files

Vor der Definition von Hochschuldaten - laut Kapitel 2.1 - wurden Files (deutsch: Dateien), wie sie üblicherweise im Download-Bereich einer Website verfügbar sind, bei der Recherche berücksichtigt. Gleichzeitig wurden *Schema.org*-Elemente gesucht, die eingebettet sind in:

```
<script type="application/ld+json">
```

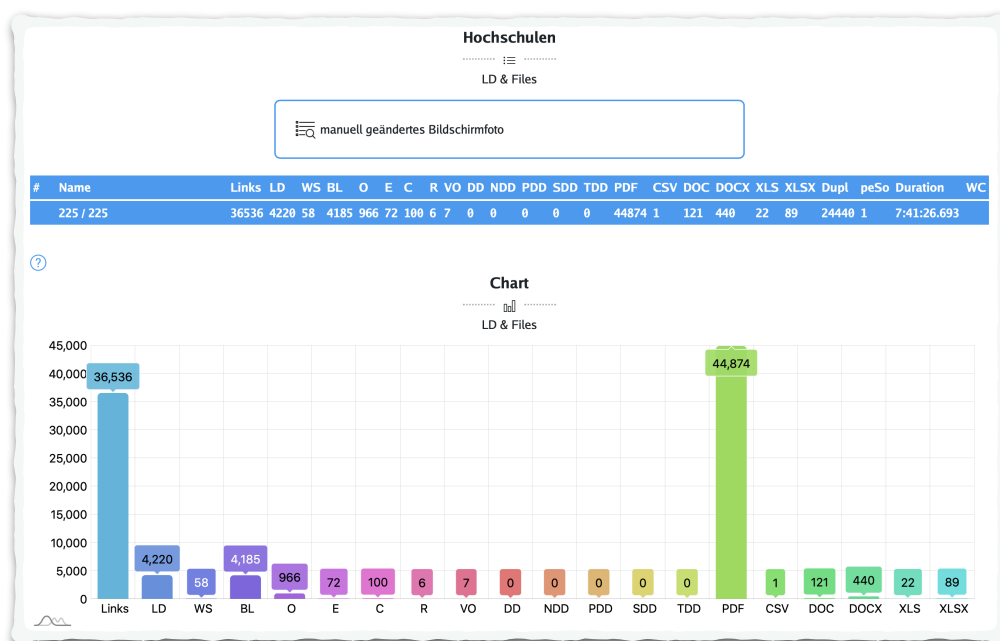


Abb. 23: LD & Files

Quelle: [Med_ind], Reiter LD & Files, Diagramm visualisiert mit amCharts v5, 27.05.2023

Abkürzungen - aus Abb. 23 - sind auf der Website via Fragezeichen-Icon abrufbar und werden ebenso im Absatz „Legende“ dieses Kapitels beschrieben.

Links

Die Summe von 36.536 setzt sich zusammen aus den Website-Aufrufen. Explizit wurde jeweils die Startseite einer Hochschuleseite angesprochen, die wiederum nach Links durchsucht wird. Eine Rekursion, sodass jeder Link erneut nach Links durchsucht wird, wurde nicht angestoßen. Grund ist, dass der Lauf - mit den gewählten Optionen in der eingesetzten Umgebung - bereits einen knappen Arbeitstag benötigt. Gleichfalls wird davon ausgegangen, dass es sich bei den zuerst erstellten Links um solche handelt, die vorwiegend im Menü integriert sind, welches sich auf jeder Unterseite kaum unterscheidet.

Legende

Weiterführende Beschreibungen, insbesondere auf die Abkürzungen in Abb. 23, können der Folgetabelle entnommen werden.

	Merkmal	Anzahl	File (Typ)	JSON-LD (Typ)
Generierung	23.04.2023			
Links	Anzahl durchsuchter Internetseiten	36.536		
LD	„Linked Data“-Elemente mit Muster: <script type="application/ld+json">	4.220		
WS	Website	58		x
BL	BreadcrumbList	4.185		x
O	Organization	966		x
E	Event	72		x
C	Course	100		x
R	Review	6		x
VO	VideoObject	7		x
DD	DigitalDocument	0		x
NDD	NoteDigitalDocument	0		x
PDD	PresentationDigitalDocument	0		x
SDD	SpreadsheetDigitalDocument	0		x
TDD	TextDigitalDocument	0		x
PDF	File-Extension	44.874	x	
CSV	File-Extension	1	x	
DOC	File-Extension	121	x	
DOCX	File-Extension	440	x	
XLS	File-Extension	22	x	
XLSX	File-Extension	89	x	
Dupl	Duplikate	24.440		
peSo	Parse Error (Beautiful-) Soup	1		
Duration	Dauer der Generierung	7h 41m		
WC	Webcode (Server-Status-Antwort)			

Tab. 20: LD & Files - Legende
Quelle: eigene Darstellung, 27.05.2023

Im Idealfall, also bei Verwendung von LD, hätten alle PDF ein DD-Äquivalent und Informationen zu XLS-/XLSX-Dateien wären in SDD-Elementen verankert.

Duplikate

Bei der Generierung ist hervorzuheben, dass die 24440 Duplikate, im Sinne eines nochmaligen Aufrufs der bereits erfassten Datei, aus Autorsicht eine stattliche Summe darstellen. Positive und negative Ausreißer werden genannt.

Domain (pos.)	Links	LD	PDF	CSV	DOC	DOCX	XLS	XLSX	Duplikate
apollon-hochschule	307	0	255	0	0	1	0	0	6
hfm-wuerzburg	92	16	138	0	0	0	0	0	3
th-owl	280	0	1.300	0	30	19	0	0	99
Domain (neg.)	Links	LD	PDF	CSV	DOC	DOCX	XLS	XLSX	Duplikate
bht-berlin	577	0	1.609	0	3	13	1	1	1.307
kh-berlin	492	0	1.035	0	0	0	0	0	2.291
th-wildau	94	0	892	0	2	13	0	3	444

Tab. 21: Duplikate - Ausreißer
Quelle: eigene Darstellung, 27.05.2023

Verhältnismäßig wenige Seiten - in Tab. 21 - besitzt die TH Wildau, stellt aber dafür eine knapp vierstellige PDF-Anzahl bereit. Dagegen ist mit drei Dateien auf der Habenseite - in Tab. 22 - die Auswahl bei der TU Clausthal sehr gering. Allerdings bietet die TU Clausthal mehr PDFs an, doch in diesem Fall erfasst der gewählte Algorithmus die fünfte Ebene [TUC_Inf] nicht.

Am Beispiel der HS Rottenburg ist in Betracht zu ziehen, dass eventuell jede Datei acht Mal verlinkt ist. Diese Aussage wurde durch einen erneuten Test zunächst dahingehend bestätigt, dass die anschließende Textsuche nach „.pdf“, in der mitgeschriebenen Ausgabe, eine Summe von 8119 lieferte. Drei weitere Duplikate endeten mit „.docx“. Prozentual ähnlich ist es bei der TU Clausthal, nur dass in diesem Fall ein PDF dreiundzwanzig Wiederholungen aufweist.

hs-rottenburg	Links	LD	PDF	CSV	DOC	DOCX	XLS	XLSX	Duplikate
23.04.2023	456	0	926	0	0	3	0	1	8.031
27.05.2023	453	0	933	0	0	3	0	1	8.122
tu-clausthal	Links	LD	PDF	CSV	DOC	DOCX	XLS	XLSX	Duplikate
23.04.2023	65	0	3	0	0	0	0	0	22
27.05.2023	66	0	3	0	0	0	0	0	24

Tab. 22: Duplikate - Vergleich
Quelle: eigene Darstellung, 27.05.2023

4.3 LD & Property

Die Konzentration liegt auf Sammlung von Eigenschaften (englisch: property), welche in diesem Kontext gleichzusetzen sind mit der Suche nach Schlüsseln (englisch: key) um ein valides JSON-LD zu erhalten. Durch Herantasten mit „randomMode“ - beschrieben in Kapitel 4.1 - wurden mehrere Property eruiert.

Über besagte Methode hinaus, beziehend auf den vorherigen Absatz, bringt der *Schema.org*-Filter [Sch_Fil] einen Mehrwert. Mittels Eingabe eines Schlagwortes (Organization, College, University) können hilfreiche Definitionen gefunden werden.

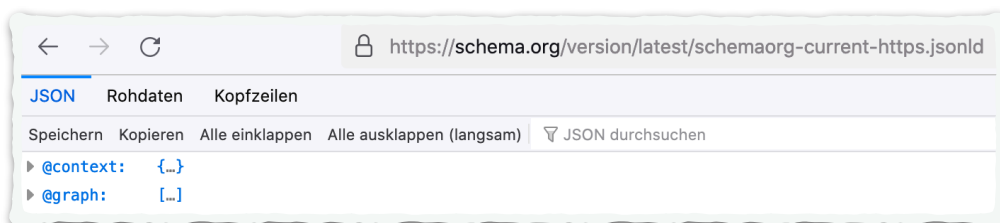


Abb. 24: Filter

Quelle: <https://schema.org/version/latest/schemaorg-current-https.jsonld>, 28.05.2023

Plausible und themenbezogene Schlüssel wurden im Eigenbau implementiert.

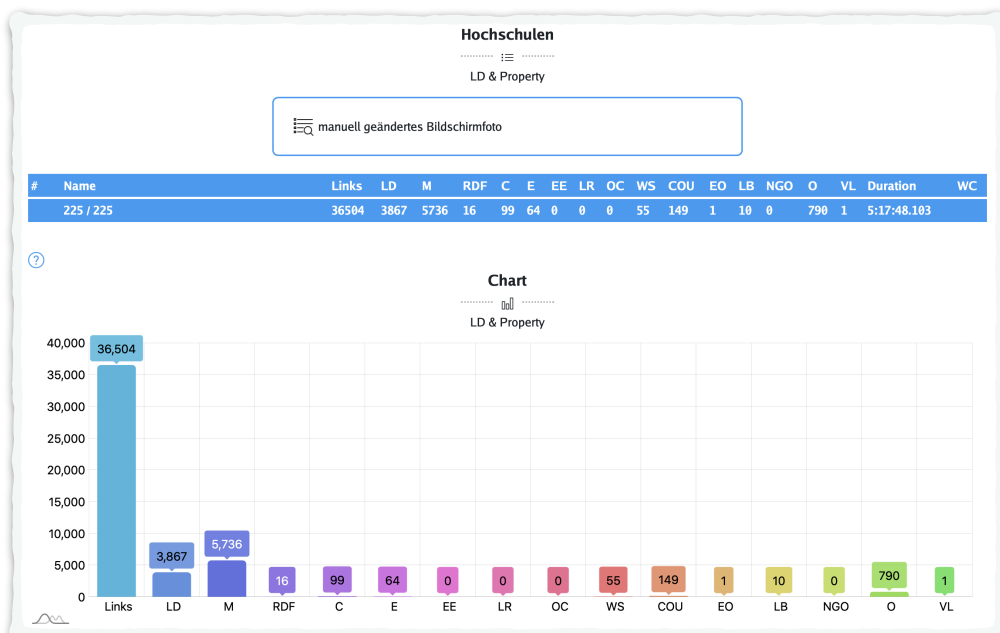


Abb. 25: LD & Property

Quelle: [Med_ind], Reiter LD & Property, Diagramm visualisiert mit amCharts v5, 28.05.2023

Legende

Für Abb. 25 werden in der Folgetabelle die Abkürzungen aufgelistet.

	Merkmal	Anzahl	File (Typ)	JSON-LD (Typ)
Generierung	28.05.2023			
Links	Anzahl durchsuchter Internetseiten	36.504		
LD	„Linked Data“-Elemente mit Muster: <script type="application/ld+json">	3.867		
M	Microdata	5.736		
RDF	Resource Description Framework	16		
C	Course	99		x
E	Event	64		x
EE	EducationEvent	0		x
LR	LearningResource	0		x
OC	OfferCatalog	0		x
WS	Website	55		x
COU	CollegeOrUniversity	149		x
EO	EducationalOrganization	1		x
LB	LocalBusiness	10		x
NGO	NonGovernmentalOrganization	0		x
O	Organization	790		x
VL	VirtualLocation	1		x
Duration	Dauer der Generierung	5h 18m		
WC	Webcode (Server-Status-Antwort)			

Tab. 23: LD & Property - Legende
Quelle: eigene Darstellung, 28.05.2023

„Links“ sind weiterhin die Aufrufe, wie bereits in Kapitel 4.2 Absatz „Links“ beschrieben. Gegenüber dem „LD & File“-Lauf liegen zweieinhalb Stunden Zeitersparnis zugrunde.

Microdata und RDF sind in dieser und allen weiteren Generierungen vertreten. Jedoch zählen sie nicht in die Typ-Statistik (Course bis WebSite, sowie CollegeOrUniversity bis VirtualLocation). Es wird nur die Anzahl ausgegeben, um Syntax-Vergleichswerte zu erhalten.

LD / M / RDF

Tab. 24 ist in drei senkrechte Bereiche geteilt, in denen jeweils die drei bestplatzierten Vertreter - bezogen auf die jeweilige Syntax - vermerkt sind. Hierbei wird die Generierung vom 28.05.2023 unter Augenschein genommen.

Domain	LD	M	RDF	Verwendung / Schema-Typ
uni-mannheim	315	0	0	- vorwiegend „BreadcrumbList“ - „Website“ (2)
uni-bamberg	307	0	0	- vorwiegend „BreadcrumbList“ und „NewsArticle“ - „Organization“ in „NewsArticle“ - „CollegeOrUniversity“
hs-kempten	211	82	1	- „BreadcrumbList“
hs-coburg	1	632	0	- vorwiegend „Person“ - „BreadcrumbList“ (33)
uol	0	538	0	- Geo-Daten der Partner-Universitäten, als Kombination „Person“ und „PostalAddress“
the-klu	0	487	0	- vorwiegend „Person“ - geringe Anzahl „PostalAddress“ und „CollegeOrUniversity“
hfmt-koeln	0	0	13	- „Zustaendigkeit“
alanus	0	1	2	- „VideoObject“
hs-kempten	211	82	1	- „BreadcrumbList“

Tab. 24: Generierungsvergleich - Syntax
Quelle: eigene Darstellung, 29.05.2023

Fügt man der Microdata-Rangliste noch die Plätze vier bis sechs hinzu (448, 440, 415) ergibt die Summe von 2960 bereits 51,6 Prozent der Gesamtanzahl für diese Syntax. Den vierten Platz belegt die TH Brandenburg und nutzt dabei vorwiegend den Schema-Typ „Person“.

Die im Archiv liegenden JS-/JSON-Dateien („until_20230603_113340“) dienen, zusätzlich zur lokalen Website, als Grundlage für Tab. 24. Diese archivierten Dateien wurden für Kontrollzwecke bei der Generierung erzeugt und haben den Vorteil, dass nicht aufgeführte Werte (Website, JSON) mitgeschrieben werden.

COU / O

Drei TYPO3-Domains sollen als Beispiel für diese Typen genutzt werden, da sie zudem zeigen, wie die Tabelle auf der lokalen Website zu handhaben ist.

Domain	Links	LD	C	COU	O
hfm-wuerzburg	84	9	0	9	0
uni-bamberg	311	307	0	15	636
wb-fernstudium	204	105	80	105	0

Tab. 25: Generierungsvergleich - Beispiele
Quelle: eigene Darstellung, 28.05.2023

Besagte Handhabung soll am Beispiel der Universität Bamberg erklärt werden.

Generell:

Es werden 311 Seiten auf der Domain untersucht, in denen 307 Muster (`<script type="application/ld+json">`) gefunden werden.

Frage:

Wieso ist „O“ (Organization) höher als „LD“ (Linked Data)?

Antwort:

Es handelt sich um den jeweiligen Typ, der gezählt wird. Die Universität Bamberg hat rund 300 Mal den Typ „NewsArticle“. In jedem „NewsArticle“ sind unter anderem die beiden Kind-Elemente „author“ und „publisher“, welche jeweils „Organization“ enthalten.

Als weiteres Beispiel wird die Wilhelm Büchner Hochschule erwähnt.

Generell:

Es werden 204 Seiten auf der Domain untersucht, in denen 105 Muster (`<script type="application/ld+json">`) gefunden werden.

Frage:

- Wieso ist die Summe von „C“ (Course) + „COU“ (CollegeOrUniversity) größer als „LD“ (Linked Data)?
- Warum sind in der JS-Datei 129 Zeilen zu finden?

Antwort:

- „COU“ ist hauptsächlich als Kind-Element in „C“ vertreten.
- Alle script-Elemente, die nicht in der Variable „scriptType“ definiert sind, werden für weitere Erkenntnisse mitgeschrieben.

4.4 Auswertung zur automatischen Recherche

Es werden nur vollständige Gesamtläufe (225 Domains) in die Auswahl genommen und auf Schlüssel reduziert, die für diesen Vergleich relevant sind. Eigenschaften weisen ein %-Zeichen auf, wenn sie in einem Lauf nicht berücksichtigt wurden. Im Folgenden ist RDF gleichzusetzen mit RDFa.

Generierung Art	Links	LD	M	RDF	C	E	WS	COU	O
23.04.2023 LD & Files	36.536	4.220	%	%	100	72	58	%	966
28.05.2023 LD & Property	36.504	3.867	5.736	16	99	64	55	149	790
11.06.2023 LD & Property	36.455	4.173	5.786	17	96	66	60	148	740
23.06.2023 LD & Property	35.976	3.535	5.299	17	19	46	53	27	152

Tab. 26: Generierungsvergleich - Auswertung
Quelle: eigene Darstellung, 23.06.2023

Bei LinkedData (LD) gibt es größere Unterschiede, die jedoch durch den Typ „BreadcrumbList“ [Med_ica] und unerreichten Domains beeinflusst werden. Rund 43% der Microdata-Werte werden durch „Person“ [ebd.] eingenommen.

Auch wenn sich die Anzahl der durchsuchten Seiten geändert hat, sind die Werte für Course (C) bis Organization (O) nahezu gleichbleibend.

Daneben werden die beiden Eigenschaften Course und CollegeOrUniversity annähernd allein von einem Fernstudium-Angebot (bis auf die technische Nichterreichbarkeit am 23.06.2023) in Anspruch genommen.

Entsprechende Vergleiche sind in Tab. 27 und Tab. 28 aufgeführt.

Domain (C)	23.04.2023	28.05.2023	11.06.2023	23.06.2023
wb-fernstudium	81	80	77	0
munich-business-school	14	14	14	14
ebz-business-school	4	4	4	4
nbs	1	1	1	1

Tab. 27: Generierungsvergleich - Course
Quelle: eigene Darstellung, 23.06.2023

Domain (COU)	23.04.2023	28.05.2023	11.06.2023	23.06.2023
wb-fernstudium	%	105	101	0
munich-business-school	%	16	16	16
uni-bamberg	%	15	18	0
hfm-wuerzburg	%	9	9	7
ebz-business-school	%	1	1	1
fh-kiel	%	1	1	1
hs-coburg	%	1	1	1
uni-leipzig	%	1	1	1

Tab. 28: Generierungsvergleich - CollegeOrUniversity
Quelle: eigene Darstellung, 23.06.2023

Durch das Ausschließen, von "application/ld+json"-Duplikaten auf der auszu-lesenden Seite, kann der Automatismus bei Bedarf optimiert werden. Es besteht die Möglichkeit die Duplikatsprüfung auf die komplette Domain zu beziehen und/oder in der JSON-Datei ein weiteres Kind-Element zu hinterlegen. Dazu-gehörige Änderungen sind im Bereich von Code 3 einzuordnen.

Des Weiteren ist ein zusätzliches Kind-Element (innerhalb von „ld“) in der JSON-Datei hilfreich, um die Anzahl der mitgeschriebenen script-Elemente (für die JS-Datei) zu hinterlegen, die nicht in der Variable „scriptType“ definiert sind.

Trotz der enormen Anzahl von durchsuchten Internetseiten (rund 36500), besteht aufgrund der gewählten Auslesetechnik (von Index alle Links nutzen), die Möglichkeit, dass Seiten - mit relevanten Daten - nicht gelesen wurden.

Erkennbar bei der Verwendung von JSON-LD ist, beziehend auf die relevanten Elemente im Sinne dieser Thesis, dass der Großteil von wenigen Domains besetzt wird. Ebenso werden Eigenschaften, wie CollegeOrUniversity, mit identischen Werten belegt und auf Unterseiten wiederholt (HfM Würzburg).

Generell bringt die automatische Recherche einen Mehrwert, da sie allein durch die Ermittlung von un-/definierten "application/ld+json"-Elementen Erkenntnisse liefert, welche Typen in welcher Anzahl auftreten. Die gleichzeitige Erstellung von Website-Bestandteilen dient zur Visualisierung der gewonnenen Daten.

5 Prototyp

Auf Grundlage der manuellen und automatischen Recherche, sowie deren Auswertungen, kann ein Prototyp im JSON-Format [Med_mid] entwickelt werden, der unter anderem auch für ein script-Tag in HTML geeignet ist.

Code 9 gibt einen Blick auf die Struktur. Testdaten der TH Brandenburg sind integriert, werden aber - besonders in „hasOfferCatalog“ - nur angerissen.

```
1 possible_LDmiddle.json > ...
2 {
3   "@context": "https://schema.org",
4   "@type": "CollegeOrUniversity",
5   "name": "TH Brandenburg",
6   "legalName": "Technische Hochschule Brandenburg",
7   "alternateName": "THB",
8   "url": "https://www.th-brandenburg.de",
9   "telephone": "+49 3381 355-0",
10  "logo": "https://www.th-brandenburg.de/typo3conf/ext/sb_config7/Resources/Public/Icons/THB_Logo.svg",
11  "photo": "https://www.th-brandenburg.de/hochschule/vorstellung-ueber-uns/campus/",
12  "slogan": "",
13  "keywords": "THB, Brandenburg, TH Brandenburg, Technische Hochschule Brandenburg, Bachelor, Master, Campus, Online, Onlinestudium",
14  "foundingDate": "1992-06-01",
15  "publicAccess": true,
16  "numberOfEmployees": [ -
17  ],
18  },
19  "sameAs": [ -
20  ],
21  },
22  "image": [ -
23  ],
24  },
25  "address": [ -
26  ],
27  },
28  "geo": [ -
29  ],
30  },
31  "hasMap": "https://www.th-brandenburg.de/kontakt-anfahrt/",
32  "openingHoursSpecification": [ -
33  ],
34  },
35  "hasOfferCatalog": [ -
36  ],
37  ],
38  "knowsLanguage": [ -
39  ]
40 }
```

Code 9: Prototyp
Quelle: [Med_mid], Visual Studio Code, 28.05.2023

Verfeinert man Abb. 11, aus Kapitel 2.1, mit JSON-Merkmalen, werden Fachbereiche und Studiengänge in „hasOfferCatalog“ inkludiert. „numberOfEmployees“ enthält die Anzahl der Professoren und Studenten. Das Resultat ist ein Mengendiagramm [Abb. 26].

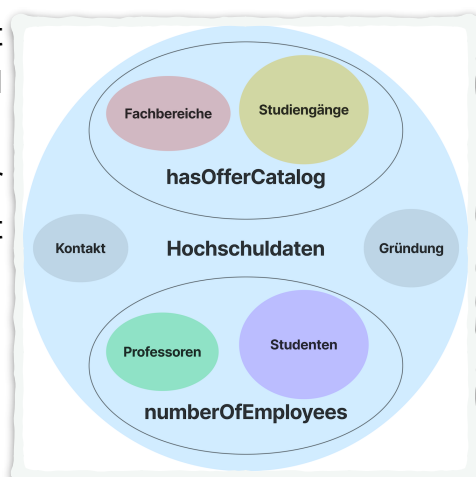


Abb. 26: Hochschuldaten - JSON
Quelle: eigene Darstellung, 28.05.2023

Allgemeine Daten

Die beiden Elemente "@context" und "@type" sind obligatorisch, damit Programme das Schema für die verwendeten Daten erkennen. Es folgen direkt aussagekräftige Angaben, die das globale Objekt (@type) beschreiben.

```
"@context": "https://schema.org",
"@type": "CollegeOrUniversity",
"name": "TH Brandenburg",
"legalName": "Technische Hochschule Brandenburg",
"alternateName": "THB",
"url": "https://www.th-brandenburg.de",
"telephone": "+49 3381 355-0",
"logo": "https://www.th-brandenburg.de/typo3conf/ext/sb_config7/Resources/Public/Icons/THB_Logo.svg",
"photo": "https://www.th-brandenburg.de/hochschule/vorstellung-ueber-uns/campus/",
"slogan": "",
"keywords": "THB, Brandenburg, TH Brandenburg, Technische Hochschule Brandenburg, Bachelor, Master, Campus, Online, Onlinestudium",
"foundingDate": "1992-06-01",
"publicAccess": true,
```

Code 10: Allgemeine Daten
Quelle: [Med_mid], Visual Studio Code, 17.06.2023

numberOfEmployees

Kind-Elemente sind als "unitText" definiert, da "unitCode" kein passendes Äquivalent besitzt [GoR_Com].

```
"numberOfEmployees": [
  {
    "unitText": "professor",
    "value": 56
  },
  {
    "unitText": "staff",
    "value": 221
  },
  {
    "unitText": "student",
    "value": 2467
  }
],
```

Code 11: numberOfEmployees
Quelle: [Med_mid], Visual Studio Code, 17.06.2023

sameAs

Link-Möglichkeiten für soziale Netze.

image

Quelle und Beschreibung eines Bildes.

```
"sameAs": [
  "https://de.wikipedia.org/wiki/Technische_Hochschule_Brandenburg",
  "https://www.facebook.com/THBrandenburg",
  "https://www.instagram.com/th_brandenburg/",
  "https://twitter.com/TH_Brandenburg",
  "https://www.youtube.com/user/FHBChannel",
  "https://de.linkedin.com/school/th-brandenburg/",
  "https://www.xing.com/net/alumni fhbrandenburg"
],
```

Code 12: sameAs
Quelle: [Med_mid], Visual Studio Code, 17.06.2023

```
"image": {
  "@type": "ImageObject",
  "url": "https://www.th-brandenburg.de/typo3conf/ext/sb_config7/Resources/Public/Icons/THB_Logo.svg",
  "caption": "Technische Hochschule Brandenburg (Logo)"
},
```

Code 13: image
Quelle: [Med_mid], Visual Studio Code, 17.06.2023

address / geo / hasMap

Ein Standort kann ebenso mit Schlüssel-Werte-Paare gekennzeichnet werden.

```
"address": {
  "@type": "PostalAddress",
  "streetAddress": "Magdeburger Straße 50",
  "addressLocality": "Brandenburg",
  "postalCode": "14770",
  "addressCountry": "DE"
},
"geo": {
  "@type": "GeoCoordinates",
  "latitude": "52.4118837",
  "longitude": "12.5367298"
},
"hasMap": "https://www.th-brandenburg.de/kontakt-anfahrt/",
```

Code 14: address - geo - hasMap
Quelle: [Med_mid], Visual Studio Code, 17.06.2023

openingHoursSpecification

Flexibel ist man bei den Öffnungszeiten.
Einzelne Tagesangaben sind möglich.

```
"openingHoursSpecification": {
  "@type": "OpeningHoursSpecification",
  "opens": "09:00",
  "closes": "17:00",
  "dayOfWeek": [
    "Monday",
    "Tuesday",
    "Wednesday",
    "Thursday",
    "Friday"
  ]
},
```

Code 15: openingHoursSpecification
Quelle: [Med_mid], Visual Studio Code, 17.06.2023

hasOfferCatalog - faculty

Fakultäten haben einen eigenen Katalog,
sowie Unterelemente für Beschreibungen.

```
"hasOfferCatalog": [
  {
    "@type": "OfferCatalog",
    "name": "faculty",
    "numberOfItems": 3,
    "itemListElement": [
      {
        "@type": "ListItem",
        "position": 1,
        "item": {
          "description": "Informatik, Applied Computer Science, Medieninformatik, Medizininformatik, Digitale Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 2,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 3,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      }
    ]
  },
  {
    "@type": "OfferCatalog",
    "name": "course of study",
    "numberOfItems": 24,
    "itemListElement": [
      {
        "@type": "ListItem",
        "position": 1,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 2,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 3,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 4,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 5,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 6,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 7,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 8,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 9,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 10,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 11,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 12,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 13,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 14,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 15,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 16,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 17,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 18,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 19,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 20,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 21,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 22,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 23,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      },
      {
        "@type": "ListItem",
        "position": 24,
        "item": {
          "description": "Informatik und Medien",
          "name": "Informatik und Medien",
          "url": "https://informatik.th-brandenburg.de/"
        }
      }
    ]
  }
],
```

Code 16: hasOfferCatalog - faculty
Quelle: [Med_mid], Visual Studio Code, 17.06.2023

hasOfferCatalog - course of study

Studiengänge belegen den zweiten Katalog in der Struktur. Damit eine Suchmaschine die Angaben zuordnen kann, sind die Listenelemente mit dem Typ "EducationalOccupationalProgram" näher definiert.

```
"hasOfferCatalog": [
  {
    "@type": "OfferCatalog",
    "name": "faculty",
    "numberOfItems": 3,
    "itemListElement": [
    ]
  },
  {
    "@type": "OfferCatalog",
    "name": "course of study",
    "numberOfItems": 24,
    "itemListElement": [
      {
        "@type": "ListItem",
        "position": 1,
        "item": {
          "@type": "EducationalOccupationalProgram",
          "name": "Medieninformatik (B.Sc.)",
          "alternateName": "OSMI",
          "description": "Die moderne Informationsgesellschaft braucht kreative Köpfe, die bei der Entwicklung innovativer M",
          "url": "https://informatik.th-brandenburg.de/studium/bachelorstudiengaenge/online-studiengang-medieninformatik/",
          "sameAs": "https://www.th-brandenburg.de",
          "timeToComplete": "P3Y",
          "numberOfCredits": 180,
          "applicationDeadline": "R/2023-09-30T22:00:00Z/P0.5Y",
          "educationalCredentialAwarded": {
            "@type": "EducationalOccupationalCredential",
            "credentialCategory": "Bachelor of Science in Computer Science"
          },
          "educationalProgramMode": [
            "online",
            "full-time",
            "part-time"
          ],
          "programPrerequisites": {
            "@type": "EducationalOccupationalCredential",
            "credentialCategory": "High school diploma"
          }
        }
      },
      {
        "@type": "ListItem",
        "position": 2,
        "item": {
          "@type": "EducationalOccupationalProgram",
          "name": "Medieninformatik (B.Sc.)",
          "alternateName": "OSMI",
          "description": "Die moderne Informationsgesellschaft braucht kreative Köpfe, die bei der Entwicklung innovativer M",
          "url": "https://informatik.th-brandenburg.de/studium/bachelorstudiengaenge/online-studiengang-medieninformatik/",
          "sameAs": "https://www.th-brandenburg.de",
          "timeToComplete": "P3Y",
          "numberOfCredits": 180,
          "applicationDeadline": "R/2023-09-30T22:00:00Z/P0.5Y",
          "educationalCredentialAwarded": {
            "@type": "EducationalOccupationalCredential",
            "credentialCategory": "Bachelor of Science in Computer Science"
          },
          "educationalProgramMode": [
            "online",
            "full-time",
            "part-time"
          ],
          "programPrerequisites": {
            "@type": "EducationalOccupationalCredential",
            "credentialCategory": "High school diploma"
          }
        }
      },
      {
        "@type": "ListItem",
        "position": 3,
        "item": {
          "@type": "EducationalOccupationalProgram",
          "name": "Medieninformatik (B.Sc.)",
          "alternateName": "OSMI",
          "description": "Die moderne Informationsgesellschaft braucht kreative Köpfe, die bei der Entwicklung innovativer M",
          "url": "https://informatik.th-brandenburg.de/studium/bachelorstudiengaenge/online-studiengang-medieninformatik/",
          "sameAs": "https://www.th-brandenburg.de",
          "timeToComplete": "P3Y",
          "numberOfCredits": 180,
          "applicationDeadline": "R/2023-09-30T22:00:00Z/P0.5Y",
          "educationalCredentialAwarded": {
            "@type": "EducationalOccupationalCredential",
            "credentialCategory": "Bachelor of Science in Computer Science"
          },
          "educationalProgramMode": [
            "online",
            "full-time",
            "part-time"
          ],
          "programPrerequisites": {
            "@type": "EducationalOccupationalCredential",
            "credentialCategory": "High school diploma"
          }
        }
      }
    ]
  }
]
```

Code 17: hasOfferCatalog - course
Quelle: [Med_mid], Visual Studio Code, 17.06.2023

Anstatt "EducationalOccupationalProgram" ist auch "Course" denkbar, doch aus Autorsicht ist die gewählte Variante eindeutiger. Prägende Merkmale sind:

- timeToComplete (Studiendauer)
- numberOfCredits (ECTS-Punkte)
- applicationDeadline (Einschreibefrist)
- educationalProgramMode (Möglichkeiten)

Um halbjährliche manuelle Aktualisierungen zu vermeiden, ist der Wert für "applicationDeadline" als Periode hinterlegt. Allerdings ist Vorsicht geboten, denn unter anderem hat die THB für internationale Studenten andere Fristen.

5.1 CollegeOrUniversity

Aufgrund einer direkteren Zuordnung und den zu vermittelnden Hochschuldaten wurde als Haupteigenschaft des Prototypen "CollegeOrUniversity" verwendet. Zudem können Elternelemente der Struktur integriert werden [Sch_COU].

Schemaspezifisch befindet man sich mit „CollegeOrUniversity“ in:

- Thing > Organization > EducationalOrganization
- Thing > Place > CivicStructure > EducationalOrganization

Schlussfolgernd ist der Prototyp mit *Schema.org* in JSON-LD beschrieben. Außerdem gilt *Schema.org* „[...] als Quasistandard für die semantische Annotation von Webseiteninhalten.“ [ION_Qua]. Die Korrektheit des Prototypen wurde durch Prüftools bestätigt. Ergänzend zum Validator #2 sei erwähnt, dass das Testtool momentan nur wenige Eigenschaften erkennt [Goo_Tes].

```
1 {
2   "@context": "https://schema.org",
3   "@type": "CollegeOrUniversity",
4   "name": "TH Brandenburg",
5   "legalName": "Technische Hochschule Brandenburg",
6   "alternateName": "THB",
7   "url": "https://www.th-brandenburg.de",
8   "telephone": "+49 3381 355-0",
9   "logo": "https://www.th-brandenburg.de/typo3conf/ext/sb_config/Resources/Public/Icons/Icons/THB_Logo.svg",
10  "photo": "https://www.th-brandenburg.de/hochschule/vorstellung-ueber-uns/campus/",
11  "slogan": "",
12  "keywords": "THB, Brandenburg, TH Brandenburg, Technische Hochschule Brandenburg, Bachelor, Master, Campus, Online, Onlinestudium",
13  "foundingDate": "1992-06-01",
14  "publicAccess": true,
15  "numberOfEmployees": [
16    {
17      "unitText": "professor",
18      "value": 56
19    },
20    {
21      "unitText": "staff",
22      "value": 221
23    },
24    {
25      "unitText": "student",
26      "value": 2467
27    }
28  ],
29 }
```

CollegeOrUniversity		0 FEHLER	0 WARNUNGEN
@type	CollegeOrUniversity		
name	TH Brandenburg		
legalName	Technische Hochschule Brandenburg		
alternateName	THB		
url	https://www.th-brandenburg.de/		
telephone	+49 3381 355-0		
logo	https://www.th-brandenburg.de/typo3conf/ext/sb_config/Resources/Public/Icons/THB_Logo.svg		
keywords	THB, Brandenburg, TH Brandenburg, Technische Hochschule Brandenburg, Bachelor, Master, Campus, Online, Onlinestudium		
foundingDate	1992-06-01		
publicAccess	http://schema.org/True		
sameAs	https://de.wikipedia.org/wiki/Technische_Hochschule_Brandenburg		
sameAs	https://www.facebook.com/THBrandenburg		
sameAs	https://www.instagram.com/th_brandenburg/		
sameAs	https://twitter.com/TH_Brandenburg		
sameAs	https://www.youtube.com/user/FHChannel		

Code 18: Validator #1
Quelle: <https://validator.schema.org>, 17.06.2023

```
1 {
2   "@context": "https://schema.org",
3   "@type": "CollegeOrUniversity",
4   "name": "TH Brandenburg",
5   "legalName": "Technische Hochschule Brandenburg",
6   "alternateName": "THB",
7   "url": "https://www.th-brandenburg.de",
8   "telephone": "+49 3381 355-0",
9   "logo": "https://www.th-brandenburg.de/typo3conf/ext/sb_config/Resources/Public/Icons/THB_Logo.svg",
10  "photo": "https://www.th-brandenburg.de/hochschule/vorstellung-ueber-uns/campus/",
11  "slogan": "",
12  "keywords": "THB, Brandenburg, TH Brandenburg, Technische Hochschule Brandenburg, Bachelor, Master, Campus, Online, Onlinestudium",
13  "foundingDate": "1992-06-01",
14  "publicAccess": true,
15  "numberOfEmployees": [
16    {
17      "unitText": "professor",
18      "value": 56
19    },
20    {
21      "unitText": "staff",
22      "value": 221
23    },
24    {
25      "unitText": "student",
26      "value": 2467
27    }
28  ],
29 }
```

Testergebnisse

2 gültige Elemente erkannt
Für gültige Elemente sind Rich-Suchergebnisse in der Google Suche möglich.
Weitere Informationen

GETESTETE SEITE ANZEIGEN VORSCHAU DER ERGEBNISSE ANZEIGEN

Erkannte strukturierte Daten

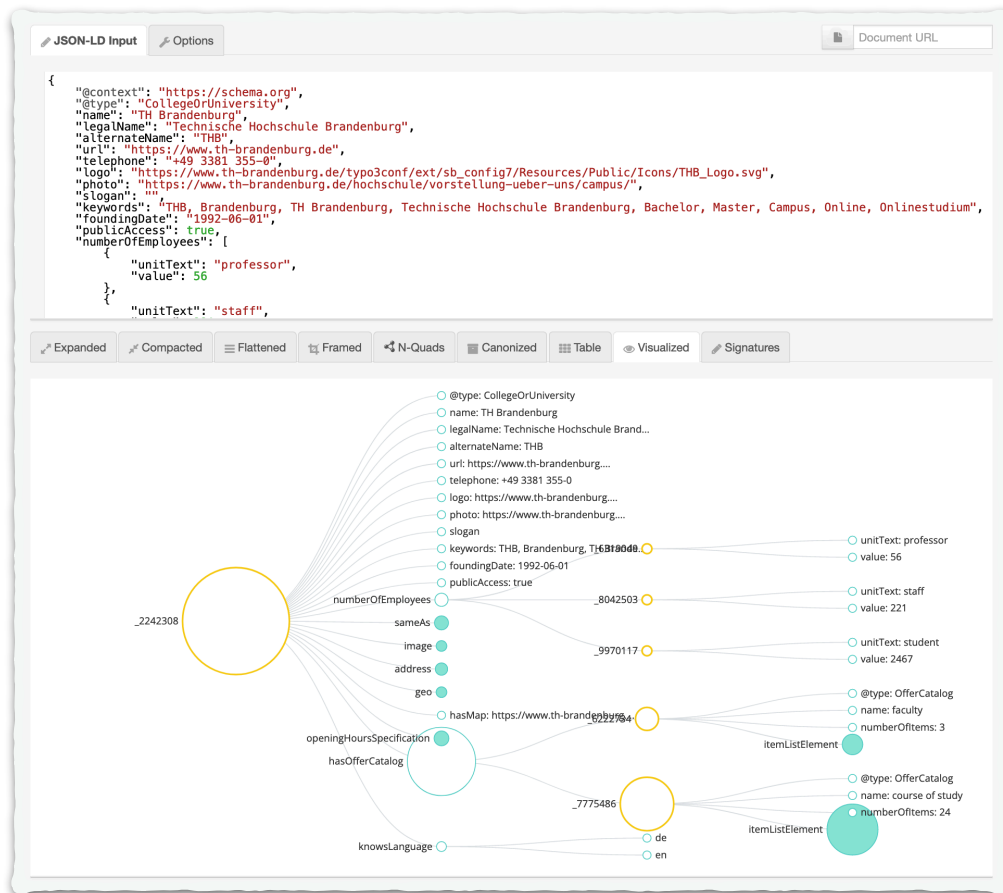
- ✓ Lokale Unternehmen 1 gültiges Element erkannt >
- ✓ Logos 1 gültiges Element erkannt >

Zusätzliche Ressourcen

- Rich-Suchergebnisse für deine gesamte Website prüfen DIE SEARCH CONSOLE AUFRUFEN

Code 19: Validator #2
Quelle: <https://search.google.com/test/rich-results>, 17.06.2023

Im Playground von *json-ld.org* kann man den Input visualisieren lassen.



Code 20: Visualized
Quelle: <https://json-ld.org/playground>, 17.06.2023

5.2 Eine Frage von TYPO3?

Kurz und prägnant: nein!

Der Prototyp basiert auf JSON und ist somit in vielen Bereichen einsetzbar. Content-Management-Systeme, wie TYPO3, können diesen auf verschiedene Arten einbinden. Einige Möglichkeiten werden auf der Folgeseite aufgelistet.

Ein weiteres Plus des Prototypen ist, dass dieser Daten enthält, die kaum gewartet werden müssen. Momentan wird ein Halbjahresrhythmus (SS/WS) prognostiziert, um die Eigenschaft "numberOfEmployees" zu aktualisieren.

5.3 Verwendungsmöglichkeiten

Generell ist der Prototyp mit wenigen Schritten auf jeder Website integrierbar.

HTML

- JSON kopieren und in den gewünschten HTML-Code einbetten
- JSON mit `<script type="application/ld+json">` und `</script>` umschließen
- JSON-Elemente mit eigenen Hochschuldaten bestücken

PHP / Datenbank

- JSON in eine Datei auslagern und einem Redakteur den Zugriff erlauben, damit die Hochschuldaten editierbar bleiben
- ein Entwickler kann mit dem „include“-Befehl diese Datei einbinden
- sollten gewünschte Hochschuldaten bereits in einer Datenbank vorliegen, kann mittels PHP - oder anderen Sprachen - das JSON erstellt werden

TYPO3

Damit der TYPO3-Redakteur in seiner gewohnten Umgebung bleibt, können zwei Extensions hilfreich sein. Mit „JSON Editor“ besteht, laut Maik Schneider von *XIMA MEDIA GmbH*, die Möglichkeit „[...] ein Input-Feld im TYPO3-Backend mit diesem Editor zu versehen, sodass Inhalte syntaktisch korrekt im JSON-Datenformat erstellt und bearbeitet werden können“. Anwendung findet JSON als Datenquelle beispielsweise in der Extension „Static Template“. TYPO3-Developer können dort sehr unkompliziert angegebene Fluid-Templates mit JSON-Daten anreichern und darstellen.

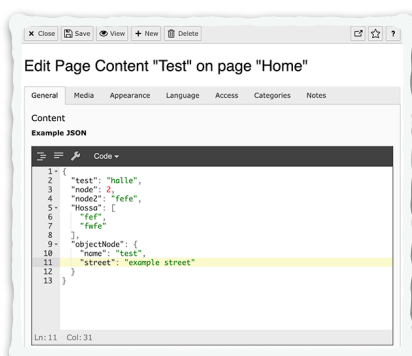


Abb. 27: JSON Editor
Quelle: <https://docs.typo3.org/p/blueways/bw-jsoneditor/main/en-us>, 22.06.2023

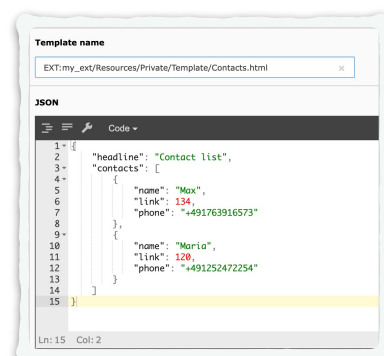


Abb. 28: Static Template
Quelle: <https://docs.typo3.org/p/blueways/bw-static-template/main/en-us>, 22.06.2023

Der Umwelt zuliebe

Nach bisheriger Einschätzung lassen sich viele Daten in JSON-LD mit *Schema.org* verpacken. So erhält ein Nutzer wertvolle Informationen und hilfreiche URLs, wenn er eine Suchmaschine in Anspruch nimmt. Aber man kann auch der Suchmaschine die Arbeit erleichtern, indem man alle relevanten JSON-LD im Code der Startseite einer Domain integriert. Daraus ergeben sich immense Vorteile:

- der Eigentümer, Entwickler oder Redakteur der Domain hat einen minimalen Organisations- und Wartungsaufwand
- man kann Unterseiten von der Indizierung ausschließen, was wiederum zur Folge hat, dass die Verwaltung von Redirects (Umleitungen) abnimmt, wenn man eine Unterseite in einem CMS archiviert oder löscht
- die Suchmaschine braucht nur die Startseite einer Domain zu durchsuchen und spart dabei Energie, Kosten und Zeit

Somit wird die - im Code verwendete - Variable „devModeOnlyIndex“, von einer Entwicklungsvariablen zwecks Zeitminimierung, zu einem wichtigen Faktor. Dabei liegt die Rechnung, unter Berücksichtigung der Recherchearten aus Kapitel 4, deutlich auf der Hand:

$$(devModeOnlyIndex = 0) \hat{=} \left(\frac{36500 \text{ Seiten}}{225 \text{ Domains}} \approx 162 \text{ Seiten pro Domain} \right) \hat{=} 18900 \text{ Sekunden}$$

$$(devModeOnlyIndex = 1) \hat{=} \left(\frac{225 \text{ Seiten}}{225 \text{ Domains}} = 1 \text{ Seite pro Domain} \right) \hat{=} 200 \text{ Sekunden}$$

Nimmt man nun die folgenden Stichpunkte an, ergibt sich eine neue Rechnung mit dem daraus resultierenden Faktor der Zeitersparnis:

- jede Domain besitzt drei Subdomains, mit je einer zu indizierenden Seite
- aus Gründen der Organisation oder Performance wurde JSON-LD aufgeteilt
- eine Sekunde pro Seite, für das Auslesen und die Verarbeitung der Daten

$$225 \text{ Domains} * \left(\frac{1 \text{ Seite}}{1 \text{ Domain}} + 3 * \frac{1 \text{ Seite}}{1 \text{ Subdomain}} \right) * \left(\frac{1 \text{ Sekunde}}{1 \text{ Seite}} \right) \hat{=} 900 \text{ Sekunden}$$

$$\text{Faktor der Zeitersparnis} = \left(\frac{18900 \text{ Sekunden}}{900 \text{ Sekunden}} \right) = 21$$

Betrachtet man den - zeitlich gesehenen - längsten Lauf (vom 23.06.2023 mit rund 24300 Sekunden) erhöht sich der Faktor der Zeitersparnis auf: 27

Professoren, Studenten, Fachbereiche und Studiengänge

Diese Attribute gehören zu den gesuchten Daten (englisch: wanted data), im Sinne dieser Thesis. Für Graphen sind Zahlen das Nonplusultra. Auf deren Grundlage lassen sich verschiedene Diagramme zeichnen. Ebenso benötigen Suchmaschinen textbasierte Informationen. Anhand des Prototypen werden beide Interessen abgedeckt.

Intelligente Programme wissen genau, welches Eltern- und Kindelement sie auslesen müssen, um an die gewünschten Daten zu gelangen. Daraus können Graphen - wie unten aufgeführt - entstehen. Um die Visualisierung zu verdeutlichen, wurden für die Abbildungen lediglich wenige Wertepaare genutzt.

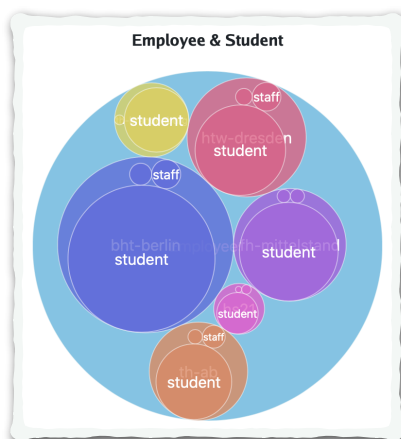


Abb. 29: Employee & Student #1
Quelle: [Med_ind], Reiter LD & Property,
Graph visualisiert mit amCharts v5, 04.06.2023

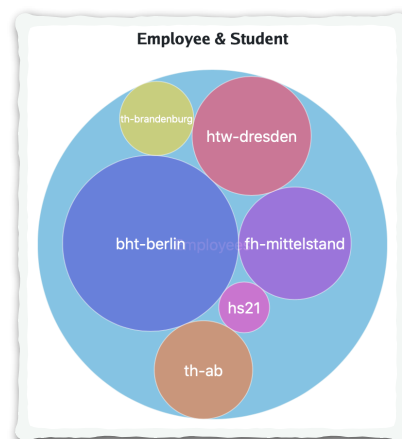


Abb. 30: Employee & Student #2
Quelle: [Med_ind], Reiter LD & Property,
Graph visualisiert mit amCharts v5, 04.06.2023

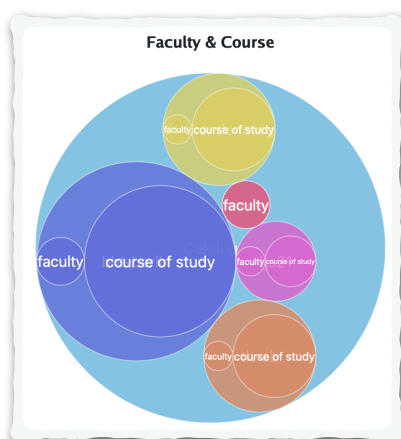


Abb. 31: Faculty & Course #1
Quelle: [Med_ind], Reiter LD & Property,
Graph visualisiert mit amCharts v5, 04.06.2023

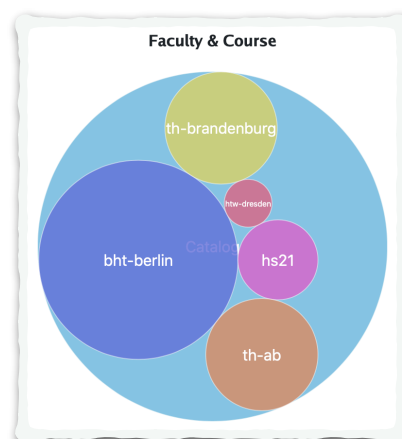


Abb. 32: Faculty & Course #2
Quelle: [Med_ind], Reiter LD & Property,
Graph visualisiert mit amCharts v5, 04.06.2023

Map

Aufgrund der Tatsache, dass im Prototyp für den Hochschulstandort ebenfalls Geo-Daten hinterlegt sind, kann eine Karte erstellt werden. Diese Idee vom Autor wurde durch die Umfrageergebnisse bestätigt. Demnach ließen sich Studenten bei ihrer Wahl der Universität oder Hochschule von der Nähe zum Wohnort beeinflussen.

Da die Karte mit Hochschuldaten erweitert wurde, kann ein Nutzer mittels „onmouseover“ (Mauszeiger auf Punkt) direkt mehr über den Standort erfahren.

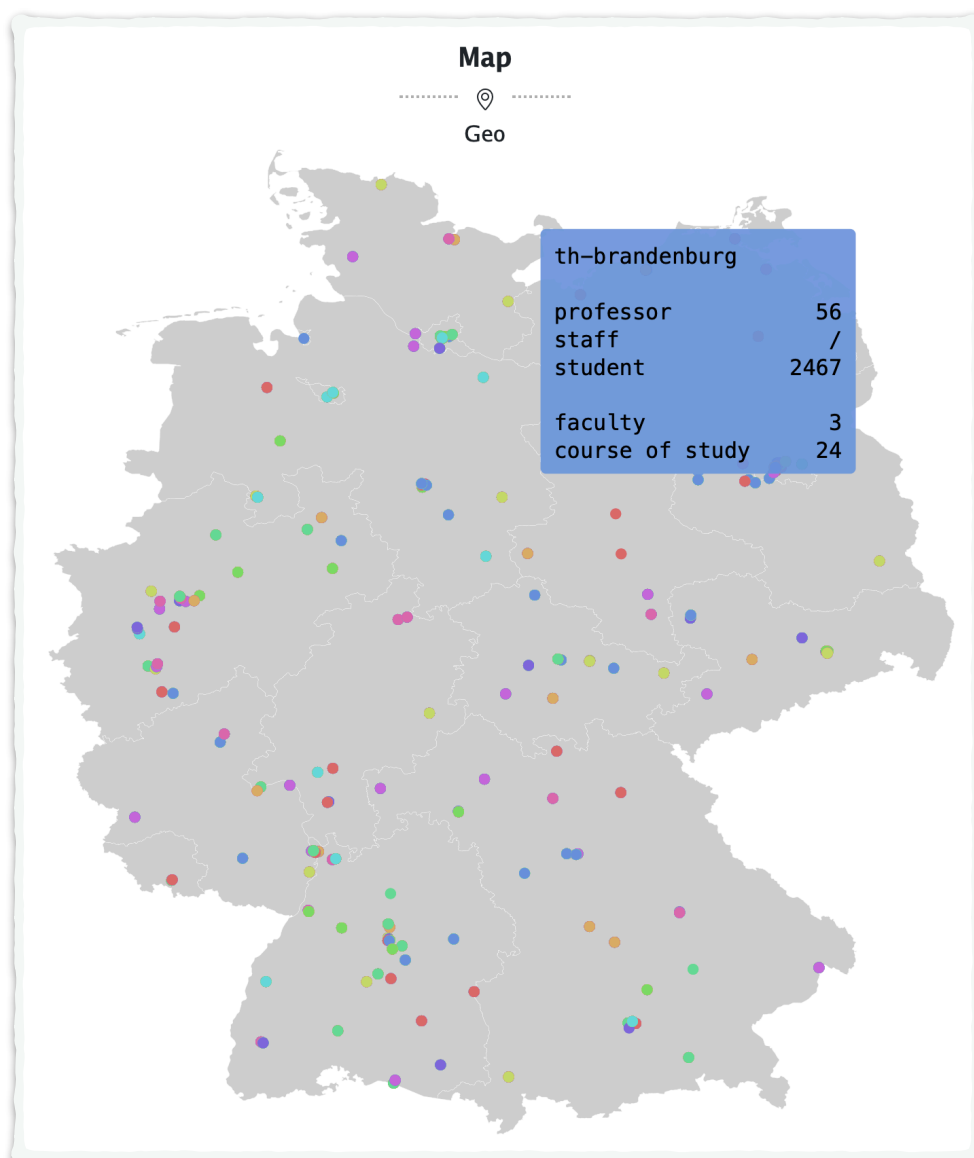


Abb. 33: Map

Quelle: [Med_ind], Reiter LD & Property, Graph visualisiert mit amCharts v5, 11.06.2023

5.4 Erweiterungsmöglichkeiten

Wie bei jedem Prototyp - oder fertigen Anwendung - ergeben sich Ansätze für Erweiterungen, um Informationen bereitzustellen:

- die Graphen (Employee & Student, Faculty & Course) für eine verbesserte Übersicht auf Bundesländer oder Großstädte reduzieren
- ein digitaler Zeitstrahl (aus den Angaben zur „Gründung“)
- die „onmouseover“-Option (Map) um „Gründung“ und „Kontakt“ erweitern
- ein Link (Klick auf Standort) in der Map ist möglich, um direkt zur Website der Universität oder Hochschule zu gelangen [Code 21]
- weitere Kontaktmöglichkeiten und Plätze hinzufügen [Code 22]

```
var container = am5.Container.new(root, {
  tooltipText: "[fontFamily: monospace]{title}[/]"/*,
  cursorOverStyle: "pointer" // Marco - 03.06.2023 - currently deactivated, but expanded tooltip with values and set fontFamily
  /*
});
container.events.on("click", (e) => {
  //window.location.href = e.target.dataItem.dataContext.url; // Marco - 03.06.2023 - currently deactivated
});
```

Code 21: Link in Map

Quelle: code/website/js/amCharts5/indexGeo.js, 22.06.2023

```
"contactPoint": [
  {
    ...
  },
  {
    "@type": "ContactPoint",
    "contactType": "student advice centre",
    "telephone": "+49 3381 355-124",
    "url": "https://www.th-brandenburg.de/studium/information-und-beratung/allgemeine-studienberatung/",
    "availableLanguage": [
      ...
    ],
    "hoursAvailable": {
      ...
    }
  },
  ...
],
"hasOfferCatalog": [
  ...
],
"knowsLanguage": [
  "de",
  "en"
],
"containsPlace": [
  {
    "name": "Hochschulbibliothek",
    "alternateName": [
      "Bibliothek",
      "Haus B"
    ],
    "address": {
      ...
    },
    "geo": {
      ...
    },
    "publicAccess": false,
    "url": "https://bibliothek.th-brandenburg.de/",
    "openingHoursSpecification": {
      ...
    }
  },
  ...
]
```

Code 22: Prototyp - Erweiterung

Quelle: [Med_pro], Visual Studio Code, 22.06.2023

6 Fazit

Als Ergebnis dieser Arbeit steht ein Prototyp im JSON-Format zur Verfügung, der mit einfachen Mitteln, nicht nur in Content-Management-Systemen, Anwendung finden kann. Die Kombination aus JSON-LD und *Schema.org* überzeugt durch Kompaktheit und einer übersichtlichen Struktur. Leichte Editierbarkeit und ein sehr geringer Wartungsaufwand sind weitere positive Faktoren.

Inhalt des Prototypen sind Hochschuldaten. Diese inkludieren Fachbereiche, Studiengänge, Kontaktmöglichkeiten und das Gründungsdatum. Hinzu kommt die Anzahl der Professoren und Studenten. Darüber hinaus finden allgemeine Informationen gleichermaßen ihren Platz.

Auf dem Weg zum Ziel wurde zunächst eine manuelle Recherche durchgeführt. Anhand deren Auswertung konnte der Fokus, in der anschließenden automatischen Recherche, auf JSON-LD gelegt werden. Hierzu dienten 225 Domains, die TYPO3 im Hintergrund verwenden, als Basis. Das Auslesen konzentrierte sich nicht nur auf die gewünschten Hochschuldaten, sondern auch auf themennahe Elemente. Zwar wurde JSON-LD verwendet, aber eher für Navigationshinweise, Events oder Personenbeschreibungen. Ein hoher Anteil, der Nichtverfügbarkeit zu implementierender Hochschuldaten, war erkennbar.

Mögliche Auswirkungen des Prototypen, wenn die darin enthaltenen Daten ausgelesen werden, wurden auf einer eigens dafür erstellten Website dargestellt. Beigemischte reale Hochschuldaten steigern das Nutzererlebnis. Elemente sind Tabellen, Graphen und eine Landkarte. Gleiche oder ähnliche Komponenten sind ebenso in einer Anwendung denkbar, die der Öffentlichkeit angeboten wird.

Bei der JSON- und Website-Erstellung konnte der Autor dieser Thesis von seiner Erfahrung, im Umgang dieser Bestandteile, profitieren. Vorhergehende Erarbeitungen diesbezüglich entfielen und wurden in Recherchen investiert. Format-Vergleiche und Quellenauswertungen waren interessant, nahmen jedoch auch Zeit in Anspruch. Letztendlich fiel die Wahl verhältnismäßig früh, aber wohlbedacht, auf JSON-LD.

Aus privater Sicht wird der Autor in naher Zukunft von JSON-LD auf seinen Seiten Gebrauch machen. Besonders der angesprochene Umweltfaktor, also nur die Startseite indizieren lassen, wurde von ihm jeher schon beachtet. Mithilfe von JSON-LD kann er weitere Daten suchmaschinengerecht anbieten.

Hochschulbezogen war das Thema ansprechend, da unterschiedliches Wissen genutzt und obendrein erweitert werden konnte. Besagtes bereits vorhandenes Know-how stammte aber größtenteils aus privaten und beruflichen Abschnitten. Der praktische Anteil ist von Vorteil, da die Theorie auf diese Art lebendig wird.

Ein Nutzen dieser Arbeit scheint von großem Interesse zu sein. Bereits vor Abgabefrist wurde der Prototyp auf TYPO3-Camps als Mundpropaganda vorgestellt. Wenn zudem entsprechende Einbindungen Anklang finden und durch die gesammelten Hochschuldaten eine wertvolle Anwendung entsteht, war die aufgebrauchte Zeit goldwert.

Das Potenzial des Prototypen ist, wie die Erweiterungsmöglichkeiten gezeigt haben, noch nicht ausgeschöpft. Die bereits genannten Hochschuldaten können in weiteren Stufen mit zusätzlichen Eigenschaften ergänzt werden. Beispielsweise sind Einschreibemodalitäten oder ein Curriculum vielgesuchte Bausteine vor einem Studium. Liegen diese als PDF vor, besteht die Möglichkeit, dass Teilangaben ebenfalls eine Rolle in JSON-LD spielen.

Durch die Implementierung des Prototypen im Wirkbetrieb, sowie die damit verbundene Bereitstellung von offenen Daten, erhalten intelligente Programme oder Suchmaschinen eindeutige und wohlgeformte Resultate. Daraus können dem Anwender nützlichere Informationen bereitgestellt werden.

Eine Win-Win-Situation für alle Beteiligten, vor allem für die Gesellschaft.

Literaturverzeichnis

- [BMI_Od] Bundesministerium des Innern und für Heimat: Open Data, in: bund, o.D., <https://www.bmi.bund.de/DE/themen/moderne-verwaltung/open-government/open-data/open-data-node.html> (abgerufen am 25.03.2023).
- [BMI_Os] Bundesministerium des Innern, für Bau und Heimat: Open-Data-Strategie der Bundesregierung, in: bund, 2021, https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/moderne-verwaltung/open-data-strategie-der-bundesregierung.pdf?__blob=publicationFile&v=5 (Download am 28.04.2023).
- [BSt_Lei] Krabina, Bernhard: Das Wichtigste auf einen Blick, in: Ein Leitfaden für offene Daten, 1. Aufl., 2020, S. 5.
- [CMS_dow] Kölbel, Judith: Universities, in: cmscensus, o.D., <https://cmscensus.eu/germany/education-culture-sport/universities> (abgerufen am 23.04.2023).
- [FHW_Sli] Fachhochschule Wedel: Startseite, in: fh-wedel, o.D., <https://www.fh-wedel.de> (abgerufen am 23.04.2023).
- [Goo_Tes] Google: Test für Rich-Suchergebnisse, in: google, o.D., <https://support.google.com/webmasters/answer/7445569#zippy=%2Cunterst%C3%BCtzte-typen> (zuletzt abgerufen am 17.06.2023).
- [GoR_Com] Good Relations: Documentation/UN/CEFACT, in: goodrelations-vocabulary, o.D., http://wiki.goodrelations-vocabulary.org/Documentation/UN/CEFACT_Common_Codes (abgerufen am 17.06.2023).
- [GoR_Cox] Good Relations: Documentation/UN/CEFACT Common Codes, Excel-Datei, in: goodrelations-vocabulary, o.D., http://wiki.goodrelations-vocabulary.org/Documentation/UN/CEFACT_Common_Codes (Download am 17.06.2023).

- [ION_Qua] IONOS: JSON-LD Tutorial nach Schema.org, in: ionos, 2018, <https://www.ionos.de/digitalguide/websites/webseiten-erstellen/tutorial-json-ld-nach-schemaorg/> (zuletzt abgerufen am 17.06.2023).
- [ION_Ver] IONOS: CMS-Vergleich 2022: Die beliebtesten Content Management Systeme, in: ionos, 2022, <https://www.ionos.de/digitalguide/hosting/cms/cms-vergleich/> (abgerufen am 23.04.2023).
- [ResGate_Kr] Kreideweiß, Sebastian: Konzeption und Implementierung einer TYPO3 Extension zur semantischen Annotation von Studiengaengen auf Web-Seiten deutscher Hochschulen, in: researchgate, 2017, <https://www.researchgate.net/project/Konzeption-und-Implementierung-einer-TYPO3-Extension-zur-semantischen-Annotation-von-Studiengaengen-auf-Web-Seiten-deutscher-Hochschulen> (Download am 25.03.2023).
- [ResGate_Me] Meister, Vera / Jonas Jetschni / Sebastian Kreideweiß: Konzept und Prototyp einer dezentralen Wissensinfrastruktur zu Hochschuldaten für Mensch und Maschine, in: researchgate, 2018, https://www.researchgate.net/publication/326732018_Konzept_und_Prototyp_einer_dezentralen_Wissensinfrastruktur_zu_Hochschuldaten_fur_Mensch_und_Maschine (Download am 25.03.2023).
- [Sch_COU] Schema.org: CollegeOrUniversity, A Schema.org Type, in: schema, 2023, <https://schema.org/CollegeOrUniversity> (zuletzt abgerufen am 17.06.2023).
- [Sch_Fil] Schema.org: Filter, in: schema, o.D., <https://schema.org/version/latest/schemaorg-current-https.jsonld> (abgerufen am 28.05.2023).
- [Sch_Rel] Schema.org: Release listing, in: schema, 2023, <https://schema.org/docs/releases.html> (zuletzt abgerufen am 17.06.2023).

- [TUC_Inf] TU Clausthal: Informatik *, in: tu-clausthal, o.D., <https://www.tu-clausthal.de/studieninteressierte/studiengaenge/bachelor-studiengaenge/informatik> (abgerufen am 25.06.2023).
- [W3C_rdf] W3C: RDF Primer, in: w3, 2014, <https://www.w3.org/TR/rdf-primer/> (abgerufen am 17.05.2023).
- [W3C_rdfa] W3C: RDFa 1.1 Primer - Third Edition, in: w3, 2015, <https://www.w3.org/TR/rdfa-primer/#validation> (abgerufen am 17.05.2023).
- [W3C_Sem] W3C: Leading the web to its full potential, in: w3, o.D., <https://www.w3.org/standards/semanticweb/> (abgerufen am 18.05.2023).
- [W3C_ttl] W3C: RDF 1.1 Turtle, in: w3, 2014, <https://www.w3.org/TR/turtle/> (abgerufen am 17.05.2023).
- [W3T_Rep] W3Techs Web Technology Surveys: Content Management Systems Market Reports, in: w3techs, o.D., https://w3techs.com/technologies/reportlist/content_management (abgerufen am 23.04.2023).
- [W3T_Usa] W3Techs Web Technology Surveys: Usage statistics and market share of TYPO3, in: w3techs, o.D., <https://w3techs.com/technologies/details/cm-typo3> (abgerufen am 25.04.2023).
- [WDC_Pro] University of Mannheim: Web Data Commons, in: webdatacommons, 2023, <https://webdatacommons.org> (abgerufen am 19.05.2023).
- [Wiki_JL] Wikimedia Foundation Inc.: JSON-LD, in: wikipedia, 2023, <https://de.wikipedia.org/wiki/JSON-LD> (abgerufen am 17.05.2023).
- [WWG_md] whatwg.org: HTML Standard, in: whatwg, o.D., <https://html.spec.whatwg.org/multipage/microdata.html#microdata> (abgerufen am 17.05.2023).

Ehrenwörtliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen oder Hilfsmittel benutzt habe und dass die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegt wurde.

Nienburg / Weser, den 6. Juli 2023

Marco Habich