

Entwicklung eines Profils für OER-Metadaten mit JSON Schema & SkoHub

Adrian Pohl

Offene Infrastruktur, Hochschulbibliothekszentrum NRW (hbz)



KIM-Workshop, WWW, 4.5.2020

Diese Präsentation:

<http://slides.lobid.org/kim-ws-2020/>



Zur Person

Angestellt am hbz, arbeite vor allem an **lobid** und in letzter Zeit an **SkoHub**

Moderator der **OER-Metadatengruppe** von KIM und Jointly

Fokus: Datenmodellierung, Web- und Metadatenstandards

Grüße an meine hbz-Kollegen (**Pascal & Fabian**) und an **Felix Ostrowski** (graphthinking GmbH). Danke für die tolle Zusammenarbeit!

Agenda

1. Was ist ein Metadatenprofil?
2. Ein LRMI-Profil für Hochschul-OER
3. Entwicklungsprozess
4. Fazit / Ausblick

Was ist ein
Metadatenprofil?

Kontext

Vorschlag von Rachel Heery & Manjula Patel im Herbst 2000, seitdem zentrales Thema in der Dublin Core Metadata Initiative (DCMI) und Verwendung in vielen anderen Kontexten

Auch "Profil", "Application Profile/Anwendungsprofil" oder "Community Profile" genannt

Aktuell Thema in der Dublin Core Application Profile Interest Group

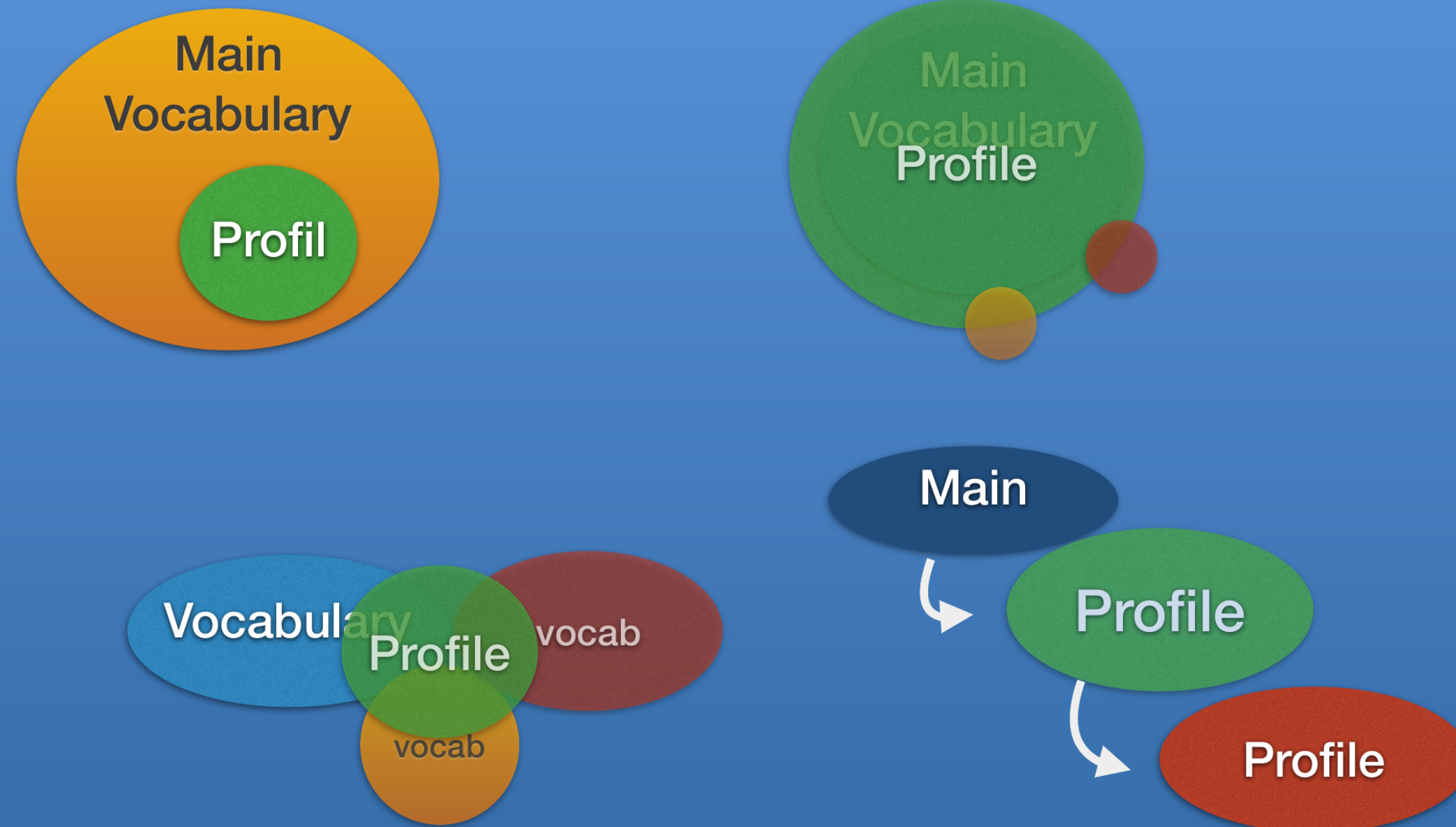
Was macht ein Profil?

Gibt an welche existierenden Metadatenschemas (z.B. Dublin Core, schema.org) und welche Elemente und Typen daraus benutzt werden

Definiert Einschränkungen für die Nutzung des/der Schema(s)

z. B. Einschränkungen auf Felder (welche Felder, Pflichtfeld: ja/nein, Häufigkeit) oder Werte (Formatvorgaben, kontrollierte Vokabulare)

What is a profile?



Quelle: Baker/Coyle (2019), Slide 4

Funktionen eines Profils

Dokumentation von Verabredungen einer Community

Zur Produktion von Daten

Für die Analyse und Validierung von Daten

Um Daten aus verschiedenen Quellen zu
mischen/auszuwählen

Um Daten Dritter zu indexieren

Für Retrieval oder das Angebot verschiedener Ansichten

Quelle: [Baker/Coyle \(2019\), Slide 3](#)

Mögliche Ausformungen/Bestandteile

Menschenlesbare Dokumentation (z.B. tabellarisch in PDFs)

Beispiel-zentrierte menschenlesbare Dokumentation (siehe
Dokumentationssession KIM WS 2017)

Schema für spezifische Serialisierung (z. B.: XML Schema,
JSON Schema)

Spezifikation eines RDF-Graphen (z. B. **ShEx**)

Generische Spezifikation eines Profils mit tabellenbasiertem
DC AP Core Vocabulary (in Entwicklung)

LRMI- Metadatenprofil für Hochschul-OER

Zu viele Akronyme?!

OER: Open Educational Resources, frei zugängliche und offen lizenzierte Lehr- und Lernmaterialien

LRMI: Learning Resource Metadata Initiative, ein Initiative zur Ergänzung von bildungsmaterialspezifischen Properties & Classes in schema.org

Wurde nicht gerade ein OER-Profil veröffentlicht?

Ja, mit **LOM for Higher Education OER Repositories** wurde bereits ein Profil inklusive XML Schema von der OER-Metadatengruppe veröffentlicht, das auf LOM (Learning Objects Metadata) basiert

LOM ist XML-basiert und eignet sich für einige Zwecke (z.B. OAI/PMH) gut, für das Web ist schema.org/LRMI besser

Warum ein LRMI-Profil?

Interesse an LRMI ist groß in der deutschsprachigen OER-Welt, allerdings gibt es keine leichte Anleitung, wie man es am besten nutzt

Aktuelles Projekt (von TIB Hannover und hbz) mit konkretem Bedarf an einem an Standards orientierten JSON Schema für OER-Metadaten: **OERSI – OER Search Index**

Der Entwicklungs- prozess

Idee zur Umsetzung

Zuerst JSON Schema & SkoHub-basiertes Referenz-
Eingabeformular entwickeln

Danach: Spezifikation mit ReSpec schreiben (siehe dazu
Sebastians Vortrag am morgigen Dienstag)

JSON Schema als zentrales Dokument

Das JSON Schema legt Felder, Wertetypen & referenzierte kontrollierte Vokabulare fest:

<https://dini-ag-kim.github.io/Irmi-profile/draft/schemas/schema.json>

Schema lässt sich im SkoHub Editor ausprobieren:

<https://skohub.io/editor/>

Kontrollierte Vokabulare als separate SKOS-Datei

Kontrollierte Vokabulare/Wertelisten werden separat als SKOS-Datei veröffentlicht und im Schema referenziert, so dass es im Editor eine Lookup-Funktion gibt

(Es gibt übrigens eine neue **deutschsprachige Einführung zu SKOS – Simple Knowledge Organization System**)

SkoHub-Editor - Mozilla Firefox

SkoHub-Editor x +

← → ↻ 🏠 🔒 https://skohub.io/editor/ ... 📄 📌 📧 📧 📧 📧 📧 📧

Skohub-Editor https://dini-ag-kim.github.io/lrmi-profile/draft/schemas/schema.json Load Schema

URL*

Title *

Type

Creator

Type *

Name*

URL

add

Description

Clear Copy to clipboard

```
{
  "@context": {
    "id": "@id",
    "type": "@type",
    "@vocab": "http://schema.org/",
    "skos": "http://www.w3.org/2004/02/skos/core#",
    "prefLabel": "skos:prefLabel",
    "inScheme": "skos:inScheme",
    "Concept": "skos:Concept"
  },
  "id": "https://example.org/kim-ws-2020-example",
  "name": "123 - Dies ist ein Titel!?",
  "type": "CreativeWork",
  "creator": [
    {
      "type": "Person",
      "name": "Hans Dampf",
      "id": "https://example.org/hansdampf"
    }
  ],
  "description": "Dies ist ein Beispiel"
}
```

Property mit nicht-formatiertem String

```
{  
  "properties": {  
    "name": {  
      "title": "Title",  
      "type": "string"  
    }  
  }  
}
```

Title *

123 - Dies ist ein Titel!?

Property mit formatiertem String (URI)

```
{
  "image": {
    "title": "Image",
    "type": "string",
    "format": "uri",
    "_display": {
      "placeholder": "A link to an image of the resource"
    }
  }
}
```

Image

A link to an image of the resource

Image

<http://dies-ist-ein-link.de>

Image

Error.format

Dies ist kein Link

Liste von möglichen Strings

```
{
  "type": {
    "title": "Type",
    "type": "string",
    "enum": [
      "AudioObject",
      "Book",
      "Course",
      "CreativeWork",
      "DataDownload",
      "ImageObject",
      "PresentationDigitalDocument",
      "SoftwareApplication",
      "VideoObject"
    ]
  }
}
```

Type

select

Creator

Type *

Name *

The creator's name

...

AudioObject

Book

Course

CreativeWork

DataDownload

Referenz auf SKOS-Vokabular

```
{
  "about":{
    "title":"Subject",
    "type":"array",
    "items":{
      "type":"object",
      "properties":{
        "inScheme":{
          "type":"object",
          "properties":{
            "id":{
              "type":"string",
              "enum":[
                "https://w3id.org/kim/hochschulfaechersystematik/"
              ]
            }
          }
        }
      }
    }
  }
}
```

Subject

H	Biblio
Lice	Geisteswissenschaften
Lice	Biblio thekswissenschaft, Dokumentation
Ima	Biblio thekswissenschaft/-wesen (nicht an Verwaltungsfachhochschulen)
htt	Rechts-, Wirtschafts- und Sozialwissenschaften
Cre	Verwaltungswissenschaften
YY	Biblio thekswesen

Wo kommen die vorgeschlagenen Werte her?

Kontrollierte Vokabulare müssen in SkoHub Vocabs veröffentlicht sein

Pflege als Turtle-Datei auf GitHub und Publikation als HTML/JSON-LD unter skohub.io (siehe [hier](#) für Details)

Hier: Nutzung einer im Kontext von "LOM for Higher Education OER Repositories" publizierten Klassifikation

hochschulfaechersystematik/hochschulfaechersystematik.ttl at master · dini-ag-kim/hochschulfaechersystematik - Mozilla Firefox

hochschulfaechersystematik x +

https://github.com/dini-ag-kim/hochschulfaechersystematik/blob/master/hochschulfaechersystematik.ttl

Search or jump to... Pull requests Issues Marketplace Explore

dini-ag-kim / hochschulfaechersystematik Unwatch 4 Star 0 Fork 2

Code Issues 1 Pull requests 1 Actions Projects 0 Wiki Security 0 Insights Settings

Branch: master hochschulfaechersystematik / hochschulfaechersystematik.ttl Find file Copy path

acka47 Update vann:preferredNamespaceUri cd087c3 on 19 Mar

2 contributors

2115 lines (1774 sloc) | 53.5 KB Raw Blame History

```
1 @base <https://w3id.org/kim/hochschulfaechersystematik/> .
2 @prefix dct: <http://purl.org/dc/terms/> .
3 @prefix skos: <http://www.w3.org/2004/02/skos/core#> .
4 @prefix schema: <http://schema.org/> .
5 @prefix vann: <http://purl.org/vocab/vann/> .
6
7 <scheme> a skos:ConceptScheme ;
8   dct:title "Destatis-Systematik der Fächergruppen, Studienbereiche und Studienfächer"@de ;
9   dct:alternative "Hochschulfächersystematik"@de ;
10  dct:description "Diese SKOS-Klassifikation basiert auf der Destatis-["Systematik der Fächergruppen, Studienbereiche und Stud
11  dct:issued "2019-12-11" ;
12  dct:publisher <https://oerworldmap.org/resource/urn:uuid:fd06253e-fe67-4910-b923-51db9d27e59f> ;
13  vann:preferredNamespacePrefix "hfs" ;
14  vann:preferredNamespaceUri "https://w3id.org/kim/hochschulfaechersystematik/" ;
15  schema:isBasedOn <http://bartoc.org/node/18919> ;
16  skos:hasTopConcept <n7>, <n1>, <n5>, <n8>, <n9>, <n4>, <n3>, <n2> .
17
18 <n1> a skos:Concept ;
19   skos:prefLabel "Geisteswissenschaften"@de ;
20   skos:narrower <n07>, <n08>, <n10>, <n13>, <n06>, <n02>, <n01>, <n09>, <n05>, <n18>, <n03>, <n14>, <n04>, <n11>, <n12>;
21   skos:notation "1" ;
22   skos:topConceptOf <scheme> .
23
24 <n2> a skos:Concept ;
25   skos:prefLabel "Sport"@de ;
26   skos:narrower <n22>;
27   skos:notation "2" ;
28   skos:topConceptOf <scheme> .
```

Mittwoch, 29. April 2020

Destatis-Systematik der Fächergruppen, Studienbereiche und Studienfächer | SkoHub-Vocabs - Mozilla Firefox

Destatis-Systematik der Fächer X +

https://skohub.io/dini-ag-kim/hochschulfaechersystematik/heads/master/w3id.org/kim/hochschulfaechersystematik/scheme.html

SkoHub-Vocabs

work this vocab on GitHub

- + 7 Agrar-, Forst- und Ernährungswissenschaften, Veterinärmedizin
- + 1 Geisteswissenschaften
- + 5 Humanmedizin/Gesundheitswissenschaften
- + 8 Ingenieurwissenschaften
- + 9 Kunst, Kunstwissenschaft
- + 4 Mathematik, Naturwissenschaften
- + 3 Rechts-, Wirtschafts- und Sozialwissenschaften
- + 2 Sport

Destatis-Systematik der Fächergruppen, Studienbereiche und Studienfächer

<https://w3id.org/kim/hochschulfaechersystematik/scheme>

Diese SKOS-Klassifikation basiert auf der Destatis-["Systematik der Fächergruppen, Studienbereiche und Studienfächer"](#).

Mittwoch, 29. April 2020

Fazit / Ausblick

Vorteile

Entwicklung auf GitHub o.ä. mit den dort etablierten
Prozessen

Referenz-Formular und Schema zur Validierung werden in
einem Zuge entwickelt

Kontrollierte Vokabulare werden parallel mit SKOS
entwickelt und sind unabhängig vom Schema nutzbar

Nachteile (?)

Fokussiert auf JSON und genau eine (von vielen möglichen) JSON-LD-Repräsentationen der Daten

Deshalb keine Validierungsmöglichkeit für anders strukturierte JSON-Daten und andere RDF-Serialisierungen

Menschenlesbare Version ist nicht von Beginn an verfügbar und muss separat geschrieben werden

Diskussion

Möglichkeiten der (zusätzlichen) Nutzung des DC Core Vocabulary for Profiles oder von ShEx?

Welche Form sollte die menschenlesbare Version haben?

Wer hat Interesse, das LRMI-Profil mit weiterzuentwickeln?

Weiterführende Informationen

- Tom Baker (2011): Dublin Core™ Application Profiles at eleven years
- Rachel Heery & Manjula Patel (2000): Application Profiles: Mixing and Matching Metadata Schemas
- SkoHub-Blogbeiträge
- Dublin Core Application Profile Interest Group
- Tom Baker, Karen Coyle (2019): Design for simple application profiles, SWIB19-Präsentation: [Slides](#), [Aufzeichnung](#)