

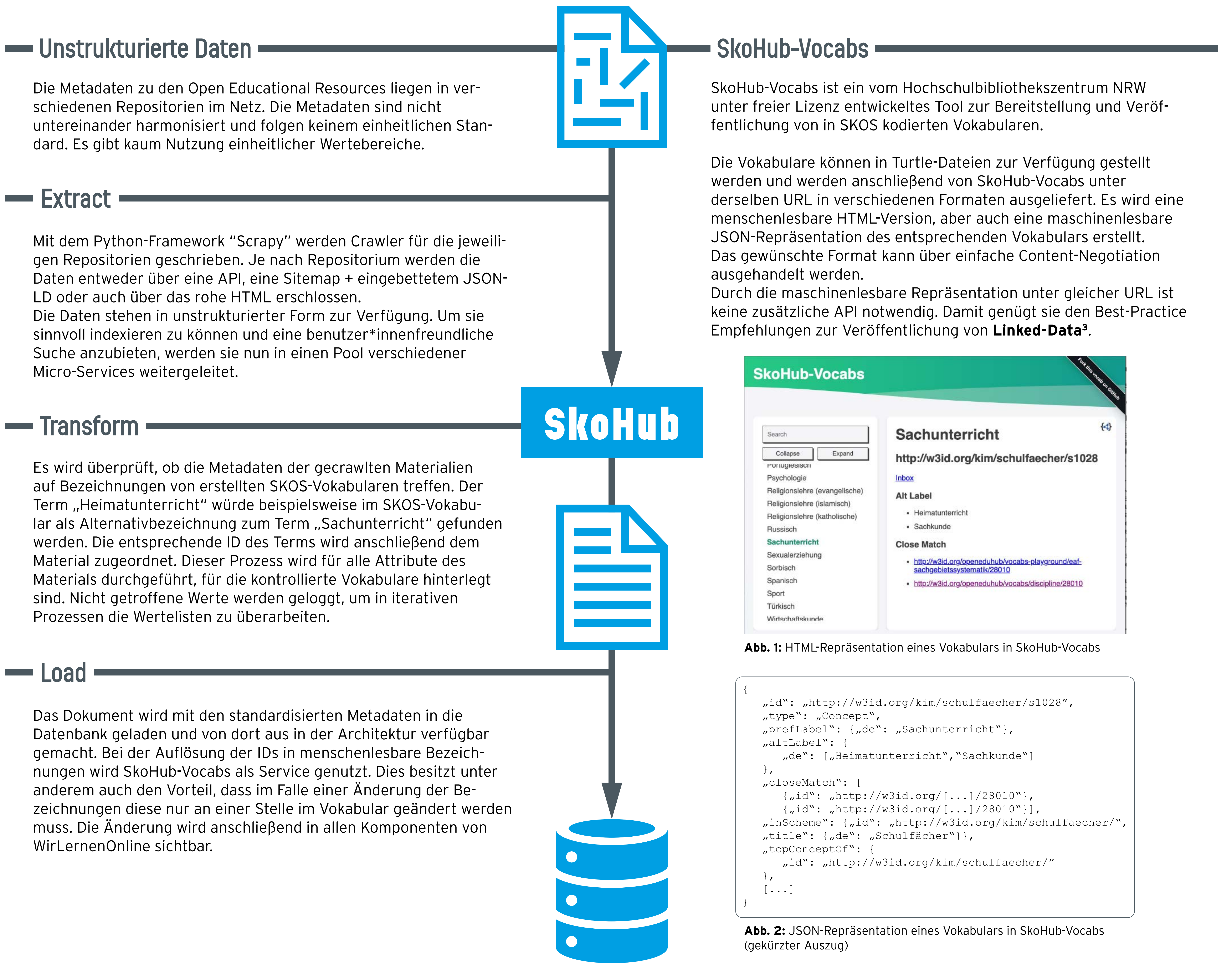
Metadatenharmonisierung in ETL-Prozessen mit SkoHub im Projekt WirLernenOnline

Steffen Rörtgen (Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen)

Abstract

Open Educational Resources (OER) sind freie Bildungsmaterialien, die unter einer offenen Lizenz veröffentlicht werden und leicht sowie rechtssicher für die eigenen Bedürfnisse angepasst werden können. Häufig liegen diese verteilt in verschiedenen Repositorien. Dort werden die Metadaten zu den Materialien meist ohne Rückgriff auf einheitliche Standards und Wertelisten gepflegt und bereitgestellt. Das Projekt „WirLernenOnline“ hat sich zu Beginn der COVID-19 Pandemie die Aufgabe gestellt, diese verteilten Materialien in einer übergreifenden Suche zentral auffindbar zu machen. Der Fokus lag dabei auf dem Bildungsbereich „Schule“. Zu diesem Zweck wurde ein ETL-Prozess entwickelt, in dem Metadaten unter Nutzung des RDF basierten Standards „SKOS“ sowie des vom Hochschulbibliothekszentrum NRW entwickelten Tools „SkoHub-Vocabs“² harmonisiert werden.

Zentral ist dabei die Idee der Trennung zwischen der Benennung und der Notation eines Terms. Statt die Werte eines Attributs als festen String zu speichern, werden lediglich IDs gespeichert, die auf SKOS-Konzepte verweisen und mittels SkoHub-Vocabs veröffentlicht werden. Alle Architekturkomponenten verwenden SkoHub-Vocabs als „Single Source of Truth“ und nutzen die dort hinterlegten Begriffe zur Anzeige im Front-End oder zur Datenharmonisierung im ETL-Prozess. Die Wertelisten wurden dabei zunächst projektspezifisch entwickelt und anschließend in der „OER-Metadatengruppe“ und „Curricula-Gruppe“ der DINI-AG-KIM generalisiert, verlinkt und wiederveröffentlicht. Von dort aus werden sie bereits in weiteren Projekten wiederverwendet, womit ein Prozess der Standardisierung und Verlinkung von Metadaten im Bildungsbereich initiiert wurde.



Ausblick: Verlinkung von Konzepten im Bildungsbereich

Die Nutzung von in SKOS kodierten Vokabularen hat bereits jetzt zu einer positiven Kettenreaktion beigetragen, indem die in der OER-Metadatengruppe und Curricula-Gruppe generalisierten Vokabulare von anderen Projekten nachgenutzt werden und sich miteinander verlinken lassen. Außerdem wird in

der Curricula-Gruppe geprüft, inwieweit sich dieser Ansatz auch auf die Abbildung und Verlinkung von Lehrplänen übertragen lässt. Hierbei ist bereits ein erster Proof-Of-Concept gelungen, der mit Hilfe des Mapping-Tools „Cocoda“⁴, entwickelt vom GBV, durchgeführt wurde.

¹ <https://wirlernenonline.de>
² <https://skohub.io>

³ <http://www.w3.org/TR/id-bp/>
⁴ <https://coli-conc.gbv.de/cocoda/>