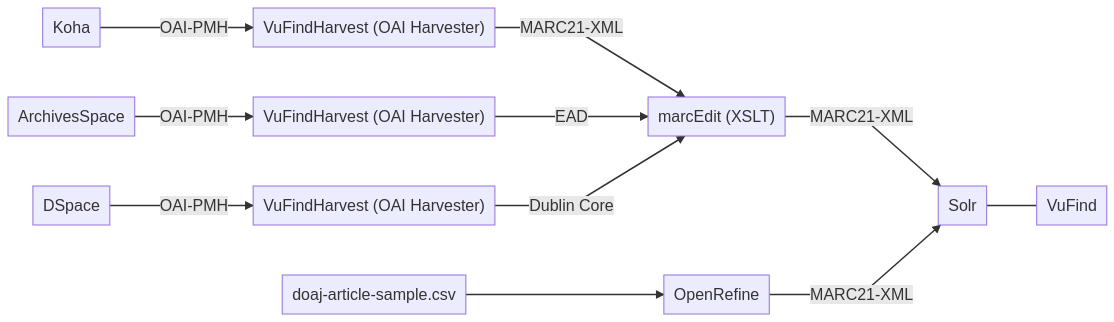
# Metadaten modellieren und Schnittstellen nutzen B



In dieser Lerneinheit fokussierten wir uns auf den Abruf der Daten und deren Konvertierung in das Marc21-Format. Dies betrifft den Bereich der Übersichtsgrafik, der bisher noch nicht bearbeitet wurde ([siehe Blog 6](https://anthonyxhd.github.io/lerntagebuch/post/6)%20Repository-Software%20für%20Publikationen%20und%20Forschungsdaten.html))



Bisher behandelt

● Koha: [Blog 4](https://anthonyxhd.github.io/lerntagebuch/post/4)%20Funktion%20und%20Aufbau%20von%20Bibliothekssystemen.html)  
● ArchiveSpace: [Blog 5](https://anthonyxhd.github.io/lerntagebuch/post/5)%20Funktion%20und%20Aufbau%20von%20Archivsystemen.html)  
● DSpace: [Blog 6](https://anthonyxhd.github.io/lerntagebuch/post/6)%20Repository-Software%20für%20Publikationen%20und%20Forschungsdaten.html)  
● OpenRefine [Blog 7](https://anthonyxhd.github.io/lerntagebuch/post/7)%20Metadaten%20modellieren%20und%20Schnittstellen%20nutzen%20A.html)

Mit [VuFind](https://en.wikipedia.org/wiki/VuFind)Harvest (OAI Harvester) kann man Daten herunterladen und dann in das MARC21-XML konvertieren. Dieser Standard wurde gewählt, weil er am weitesten verbreitet ist und bei der Konvertierung immer mit Verlusten gerechnet werden muss.

## Austauschprotokolle für Metadaten

Von den zahlreichen Übertragungsprotokollen im Bibliotheks- und Archivbereich sind drei besonders weit verbreitet und relevant:

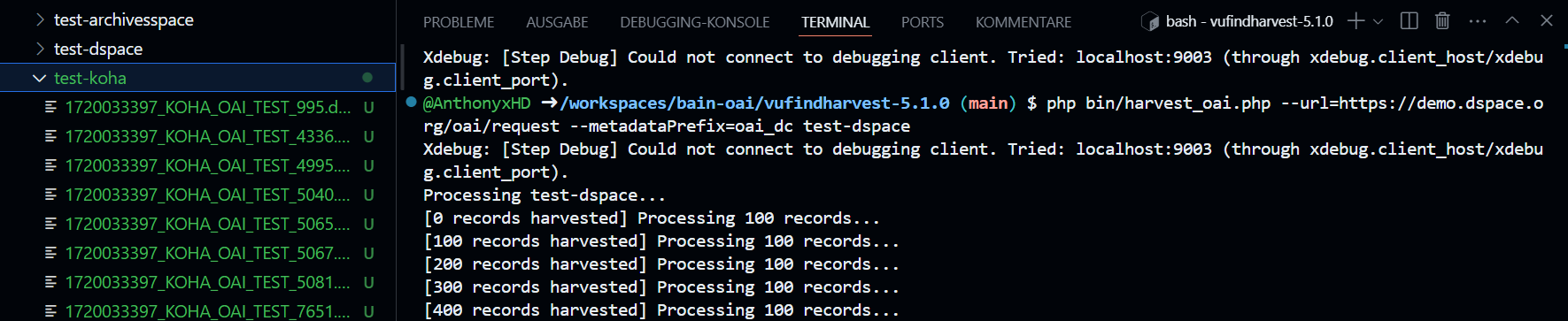
● [Z39.50](https://www.loc.gov/z3950/agency/)  
● [SRU](https://www.loc.gov/standards/sru/)  
● [OAI-PMH](https://www.openarchives.org/pmh/)

Z39.50 und SRU eignen sich besonders für Live-Abfragen oder gezielte Datenabrufe mit vielen Parametern, während OAI-PMH vor allem für größere Datenabzüge und regelmäßige Aktualisierungen geeignet ist. Zu beachten ist, dass Z39.50 zwar sehr alt ist, aber immer noch im Einsatz. Die Arbeit mit diesem Protokoll kann nicht direkt über den Browser erfolgen und benötigt daher Zusatzsoftware.

## Metadaten über OAI-PMH harvesten mit VuFindHarvest

Um die Übungen zu lösen, haben wir erneut mit einem Codespace auf GitHub gearbeitet ([Blog 7](https://anthonyxhd.github.io/lerntagebuch/post/7)%20Metadaten%20modellieren%20und%20Schnittstellen%20nutzen%20A.html)). Zu Beginn luden wir die Datensätze aus den entsprechenden Demo-Installationen herunter. Dazu mussten wir die Prefixes finden, um das Kommando zum Aufruf des Harvestings anzupassen. Dies führte dazu, dass wir die Daten im richtigen Format herunterluden:

● Koha: –metadataPrefix=marc21  
● ArchivesSpace: –metadataPrefix=oai\_ead  
● DSpace: –metadataPrefix=oai\_dc

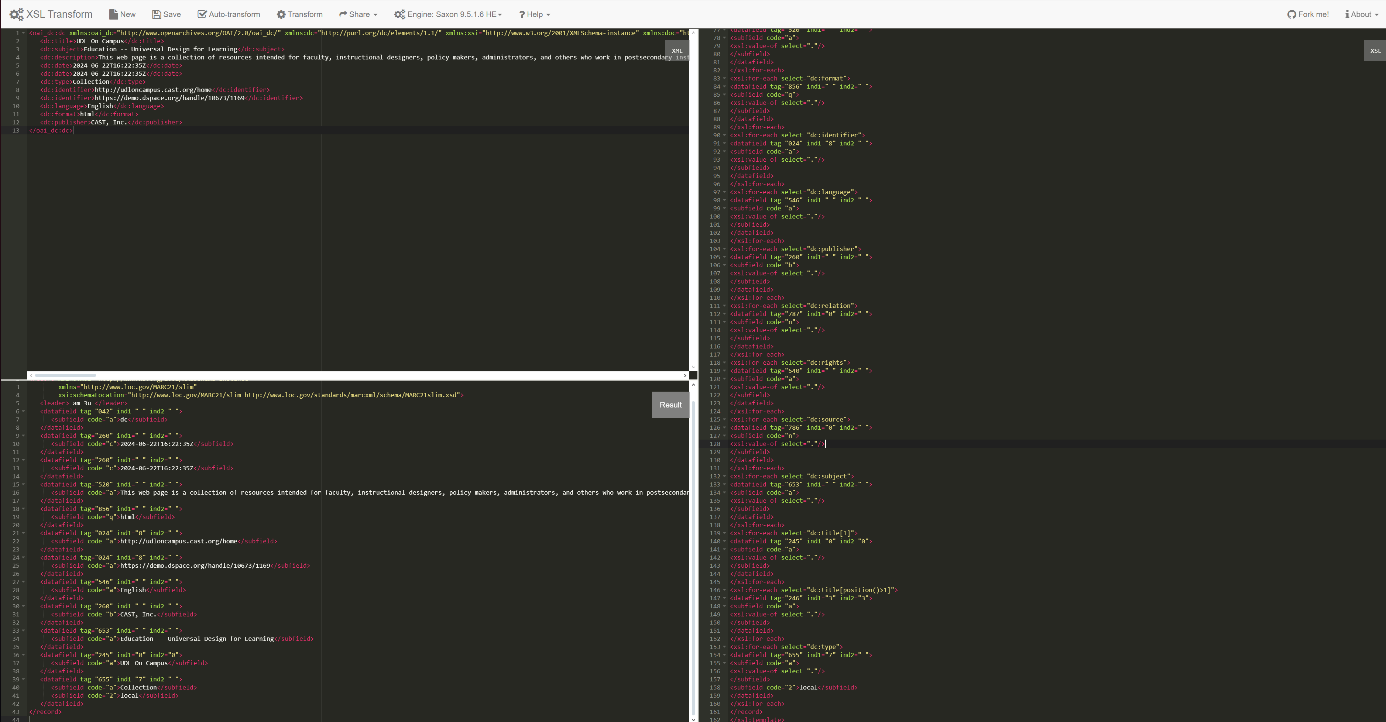


## Konvertierung / Crosswalk

Crosswalks sind ein gängiger Begriff, um die Konvertierung von einem Metadatenstandard in einen anderen zu beschreiben ([coole Übersicht](https://www.linkedin.com/advice/0/how-can-you-design-effective-metadata-crosswalks-fk9gc)). Ein Beispiel dafür ist die Umwandlung von Dublin Core zu MARC21. Der "Crosswalk" beinhaltet Regeln, wie Elemente und Werte zugeordnet werden, auch bekannt als Mapping. Obwohl dies im Idealfall verlustfrei erfolgen soll, ist meist keine 1:1-Zuordnung möglich.

[XSLT (Extensible Stylesheet Language Transformations)](https://en.wikipedia.org/wiki/XSLT#:~:text=XSLT%20(Extensible%20Stylesheet%20Language%20Transformations,as%20PDF%2C%20PostScript%20and%20PNG.) ist eine Programmiersprache zur Transformation von XML-Dokumenten. Sie wurde 1999 als W3C-Empfehlung veröffentlicht und war ursprünglich dafür gedacht, ein XML-Dokument in ein anderes zu transformieren.

Versuch eine heruntergeladene Datei zu konvertieren mit dem [Online-Tool](http://xsltransform.net)



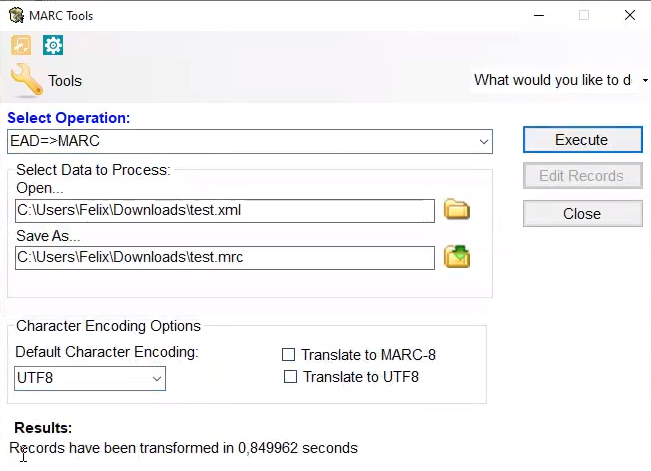
Beispiele der Library of Congress (LoC)

● [MARC21 zu DC](https://www.loc.gov/marc/marc2dc.html)  
● [DC zu MARC21](https://www.loc.gov/marc/dccross.html)

Ziel des Moduls ist es nicht, ein XSLT-Profi zu werden, sondern den Konvertierungsprozess zu durchschauen und zu verstehen. Hierzu ist das erste Beispiel der Library of Congress sehr hilfreich und verständlich.

## MarcEdit

Am Ende der Lernveranstaltung wurde die kostenlose, aber nicht Open Source-Software "MarcEdit" vorgeführt, die häufig bei der Arbeit mit MARC21 verwendet wird.



Die Schritte, wie man in solchen Fällen vorgeht, sind [hier](https://pad.gwdg.de/vv-Q4RDnQ22Gqc-EEgbewA#Transformation-von-ArchivesSpace-EAD-zu-MARC21-Versuch-2) beschrieben

## Learnings

In dieser Lerneinheit lernte ich, wie man Daten herunterlädt und in das MARC21-Format konvertiert. Ich erlangte ein Verständnis der wichtigsten Metadatenübertragungsprotokolle wie Z39.50, SRU und OAI-PMH sowie deren Einsatzgebiete.

In einer Übung setzten wir VuFindHarvest zur Datenernte ein. Dieser Teil gefiel mir besonders gut, da der Praxisbezug und die Arbeit mit der Shell das Ganze für mich interessant machten.

Zuletzt wurden mir Kenntnisse über Crosswalks und die Herausforderungen der verlustfreien Konvertierung zwischen verschiedenen Metadatenstandards vermittelt. XSLT wurde ebenfalls gezeigt. Das Grundprinzip habe ich zwar verstanden, aber eine vertiefte Arbeit wäre hierzu notwendig. Das Interesse ist jedoch nicht groß genug, um die Zusatzübungen zu machen.