

# KIM Workshop 2021

<b>Wann:</b>	26. bis 28. April 2021
<b>Wo:</b>	online
<b>Veranstalter:</b>	DINI-AG KIM, UB Mannheim
<b>Programmkomitee+Organisationsteam:</b>	Tracy Arndt (DNB), Alex Jahnke (SUB Göttingen), Jana Hentschke (ZBW Hamburg), Sylvia Hulin (UB Mannheim), Stefanie Rühle (SUB Göttingen), Jakob Voß (GBV), Philipp Zumstein (UB Mannheim)
<b>Twitter:</b>	Hashtag <a href="#">#kimws21</a>
<b>Kontakt:</b>	<a href="mailto:kim-info@din.de">kim-info@din.de</a>


## Programm

[ [Montag, 26. April 2021](#) ] [ [Dienstag, 27. April 2021](#) ] [ [Mittwoch, 28. April 2021](#) ]

### Montag, 26. April 2021

0  
9  
:  
3  
0  
-  
1  
3  
:  
00

Hands-On Tutorials

 Anmeldung erforderlich

Option 1: Einführung in die XML Datenbank eXist-db	Option 2: Von 0 auf Triple Store in 180 Minuten
<p><b>Leitung:</b> Uwe Sikora (SUB Göttingen)</p> <p>In diesem Hands-On wird in die grundlegenden Features und Konzepte der XML-Datenbank eXist-DB eingeführt. Zunächst werden wir uns anschauen, wie man eine eXist Instanz konfiguriert und managt.</p> <p>Im Anschluss verschaffen wir uns einen Überblick über die Struktur von eXist und seiner eigenen Entwicklungsumgebung "eXide". In diesem Rahmen werden die beiden Basiskonzepte - APPs und Libraries - vorgestellt, mit denen man für eXist entwickelt: Wir werden eine eigene eXist Anwendung erstellen und sie mit XML Daten befüllen.</p> <p>Darauf aufbauend widmen wir uns der Abfrage von gespeicherten Daten via XPath und XQuery: Nach einer praktischen Einführung in XPath &amp; XQuery nehmen wir das Abfragekonzept von eXist etwas detaillierter unter die Lupe und schreiben eigene Queries. In diesem Rahmen werden wir uns auch mit den Indexierungsmöglichkeiten beschäftigen, die eXist bietet und eigene Indexkonfigurationen erstellen, die wir in einer einfachen Keyword-In-Context Suche implementieren.</p> <p>Um das Bild abzurunden, betrachten wir in einem Ausblick die RESTXQ Schnittstelle, mit deren Hilfe RESTfull Services in eXist implementiert werden können.</p> <p><b>Zielgruppe:</b> XML Entwickler und alle, die mit XML Daten arbeiten</p> <p><b>Vorkenntnisse:</b> XML, XPath (wünschenswert), XQuery (wünschenswert)</p> <p><b>Voraussetzungen:</b> Eine lokal installierte eXist-db 5.2.0 Instanz (<a href="http://exist-db.org/exist/apps/homepage/index.html">http://exist-db.org/exist/apps/homepage/index.html</a>)</p>	<p><b>Leitung:</b> Andreas Walker (SUB Göttingen)</p> <p>Wie baue ich eigentlich meinen eigenen Triple Store auf? In nur 180 Minuten soll dieses Hands-On-Tutorial diese grundlegende Frage beantworten. In einem ersten Teil wollen wir zunächst anhand eines existierenden Triple Stores (z.B. Wikidata) lernen, wie Daten in einem Triple Store strukturiert sind, und wie man sie mit Hilfe der Sprache SPARQL abfragen kann. Nach der Pause nutzen wir diese neuen Kenntnisse dann, um unsere eigenen Daten als RDF/Turtle aufzubereiten und in einem Triple Store zu hosten. Wir setzen dabei auf die Open Source-Software <a href="#">Apache Fuseki</a>. Ziel ist es, die TeilnehmerInnen über die erste Hürde des Einstiegs ins Semantic Web zu bringen, damit sie selbstständig weiterdenken und -experimentieren können.</p> <p><b>Zielgruppe:</b> Personen mit Interesse am Semantic Web und Spaß am Experimentieren</p> <p><b>Vorkenntnisse:</b> Interesse am Umgang mit Daten und Datenbanken. Vorkenntnisse mit SPARQL sind sehr willkommen, aber nicht notwendig.</p> <p><b>Voraussetzungen:</b> Für den ersten Teil genügt ein Notebook mit Internetzugang und modernem Webbrowser. Im zweiten Teil sollten die TeilnehmerInnen idealerweise über Adminrechte auf ihrem Rechner verfügen und mindestens Java 8 installiert haben, um eine eigene Fuseki-Instanz betreiben zu können. Wenn das nicht möglich ist, versuchen wir aber trotzdem Lösungen für alle zu finden (z.B. Arbeit in Gruppen, Fuseki auf externen Servern). Eine kurze Vorab-E-Mail in diesem Fall an <a href="mailto:walker@sub.uni-goettingen.de">walker@sub.uni-goettingen.de</a> spart uns Zeit im Workshop</p>

1 3 : 0 0 - 1 4 : 00	Programmpause
1 4 : 0 0 - 1 4 : 10	<b>Begrüßung</b>  Alex Jahnke (SUB Göttingen), Jana Hentschke (ZBW Hamburg/Kiel) (DINI-AG KIM AG-Sprecher und -Sprecherin)
1 4 : 1 0 - 1 4 : 40	Vortrag  <b>Paul Vierkant (datacite): Datenkuratierung, Entwicklung eines Metadatenschemas und Abfragemöglichkeiten für ROR-IDs (Folien <a href="#">online</a> und direkt als <a href="#">PDF</a>)</b>  Das Research Organization Registry (ROR) hat sich als offener Persistent Identifier (PID) für wissenschaftliche Einrichtungen auf den Weg gemacht ein globaler Standard zu werden. DataCite, California Digital Library und Crossref stellen diesen 2019 gestarteten Open Science Service bereit. Dieser Beitrag stellt ROR, sein Metadatenschema, seine API-Funktionen und den Ansatz der „community-based“ Datenkuratierung vor. Zudem werden Einblicke in die Ergebnisse der jüngsten Umfrage zu Organisations-IDs in Deutschland im Rahmen des Projekts ORCID DE gegeben. Feedback im Rahmen des Vortrags zu ROR ist gern gesehen.
1 4 : 4 0 - 5 : 10	Vortrag  <b>Thomas Baker (DCMI): Jenseits von Simple und Qualified – Dublin Core heute (<a href="#">Folien</a>)</b>  "Metadaten" hiess immer Datenbanken oder Formate - massgeschneiderte Strukturen für geschlossene Systeme. Die wilde Vielfalt des offenen WWW führte jedoch zu Linked Data: Metadaten als Bündel einzelner Aussagen, die auf der Basis bekannter Vokabulare und generischer Datenmodelle rekombinierbar sind.  Die älteren und neueren Sichtweisen bestehen nebeneinander noch im Anwenderkreis um Dublin Core, werden jedoch von der Idee des Anwenderprofils überbrückt, die ihrerseits von neuen Techniken für die Validierung offener Daten (ShEx und SHACL) verstärkt wird.  <a href="#">Folien mit Notizen</a>
1 5 : 1 0 - 1 5 : 30	Programmpause
1 5 : 3 0 - 1 7 : 00	Interaktive Session  <b>Metadaten in der Lehre</b>  <b>Moderation:</b> Tracy Arndt (DNB)  In dieser partizipativen Session wollen wir über die Inhalte und den Stellenwert von Metadaten in den einschlägigen deutschsprachigen Studiengängen diskutieren. Darüber hinaus wollen wir über die Anforderungen aus der Praxis im Bereich des Metadatenmanagements sprechen. Welche für den Berufsalltag notwendigen Kompetenzen werden bereits in der Lehre vermittelt und welche scheinen noch offen? Sollten bestimmte Formate, Tools und theoretische Betrachtungen vermittelt werden?

Dienstag, 27. April 2021

<b>Option 1:</b>  <b>Von 0 auf Triple Store in 180 Minuten</b>	<b>Option 2:</b>  <b>Processing MARC</b>	<b>Option 3:</b>  <b>Einführung in die XML Datenbank eXist-db</b>
<p><b>Leitung:</b> Andreas Walker (SUB Göttingen)</p> <p>[Wiederholung des Tutorials vom Vortag]</p> <p>Wie baue ich eigentlich meinen eigenen Triple Store auf? In nur 180 Minuten soll dieses Hands-On-Tutorial diese grundlegende Frage beantworten. In einem ersten Teil wollen wir zunächst anhand eines existierenden Triple Stores (z.B. Wikidata) lernen, wie Daten in einem Triple Store strukturiert sind, und wie man sie mit Hilfe der Sprache SPARQL abfragen kann. Nach der Pause nutzen wir diese neuen Kenntnisse dann, um unsere eigenen Daten als RDF/Turtle aufzubereiten und in einem Triple Store zu hosten. Wir setzen dabei auf die Open Source-Software <i>Apache Fuseki</i>. Ziel ist es, die TeilnehmerInnen über die erste Hürde des Einstiegs ins Semantic Web zu bringen, damit sie selbstständig weiterdenken und -experimentieren können.</p> <p><b>Zielgruppe:</b> Personen mit Interesse am Semantic Web und Spaß am Experimentieren</p> <p><b>Vorkenntnisse:</b> Interesse am Umgang mit Daten und Datenbanken. Vorkenntnisse mit SPARQL sind sehr willkommen, aber nicht notwendig.</p> <p><b>Voraussetzungen:</b> Für den ersten Teil genügt ein Notebook mit Internetzugang und modernem Webbrowser. Im zweiten Teil sollten die TeilnehmerInnen idealerweise über Adminrechte auf ihrem Rechner verfügen und mindestens Java 8 installiert haben, um eine eigene Fuseki-Instanz betreiben zu können. Wenn das nicht möglich ist, versuchen wir aber trotzdem Lösungen für alle zu finden (z.B. Arbeit in Gruppen, Fuseki auf externen Servern). Eine kurze Vorab-E-Mail in diesem Fall an <a href="mailto:walker@sub.uni-goettingen.de">walker@sub.uni-goettingen.de</a> spart uns Zeit im Workshop</p>	<p><b>Leitung:</b> Johann Rolschewski (SBB Berlin)</p> <p>2002 erklärte <a href="#">Roy Tennant "MARC Must Die"</a>. Aktuell ist das <a href="#">Format "MARC 21"</a> immer noch das meist genutzte Format zum Austausch und Katalogisierung von Metadaten in Bibliotheken. Selbst unsere "Next Generation Library Systems" verwenden diesen Standard aus den 60er Jahren. Da wir auch in den kommenden Jahren mit "MARC 21" arbeiten werden, soll dieses Tutorial eine Einführung zu folgenden Themen geben:</p> <ul style="list-style-type: none"> <li>■ Struktur von "MARC 21"-Datensätzen und ihre verschiedenen Serialisierungen (MARCXML, MARCmaker, MARC-in-JSON, ALEPHSEQ)</li> <li>■ Validierung von "MARC 21"-Datensätzen und häufige Fehler</li> <li>■ Statistische Auswertung von "MARC 21"-Datensätzen</li> <li>■ Konvertierung von "MARC 21"-Datensätzen</li> <li>■ Metadatenextraktion aus "MARC 21"-Datensätzen</li> </ul> <p>Das Tutorial richtet sich an Systembibliothekarinnen und Datenmanagerinnen. Für die meisten Aufgaben werden wir Kommandozeilen-Tools wie <i>yaz-marcdump</i>, <i>marcstats</i>, <i>marcvalidate</i> und <i>catmandu</i> verwenden, daher sollten die TeilnehmerInnen mit den Grundlagen der Kommandozeile (CLI) vertraut sein. Für die Übungen wird ein Rechner mit <a href="#">installierter VirtualBox</a> benötigt. Der Veranstalter stellt im Vorfeld eine <a href="#">"Virtual Machine" (VM, Linux-Gastsystem)</a> zur Verfügung. Die TeilnehmerInnen können ihre eigenen "MARC 21"-Datensätze für die praktischen Übungen mitbringen.</p> <p><b>Link zum Handbuch:</b> <a href="https://jorol.github.io/processing-marc/#/">https://jorol.github.io/processing-marc/#/</a></p>	<p><b>Leitung:</b> Uwe Sikora (SUB Göttingen)</p> <p>[Wiederholung des Tutorials vom Vortag]</p> <p>In diesem Hands-On wird in die grundlegenden Features und Konzepte der XML-Datenbank eXist-DB eingeführt. Zunächst werden wir uns anschauen, wie man eine eXist Instanz konfiguriert und managt.</p> <p>Im Anschluss verschaffen wir uns einen Überblick über die Struktur von eXist und seiner eigenen Entwicklungsumgebung "eXide". In diesem Rahmen werden die beiden Basiskonzepte - APPs und Libraries - vorgestellt, mit denen man für eXist entwickelt: Wir werden eine eigene eXist Anwendung erstellen und sie mit XML Daten befüllen.</p> <p>Darauf aufbauend widmen wir uns der Abfrage von gespeicherten Daten via XPath und XQuery: Nach einer praktischen Einführung in XPath &amp; XQuery nehmen wir das Abfragekonzept von eXist etwas detaillierter unter die Lupe und schreiben eigene Queries. In diesem Rahmen werden wir uns auch mit den Indexierungsmöglichkeiten beschäftigen, die eXist bietet und eigene Indexkonfigurationen erstellen, die wir in einer einfachen Keyword-In-Context Suche implementieren.</p> <p>Um das Bild abzurunden, betrachten wir in einem Ausblick die RESTXQ Schnittstelle, mit deren Hilfe RESTfull Services in eXist implementiert werden können.</p> <p><b>Zielgruppe:</b> XML Entwickler und alle, die mit XML Daten arbeiten</p> <p><b>Vorkenntnisse:</b> XML, XPath (wünschenswert), XQuery (wünschenswert)</p> <p><b>Voraussetzungen:</b> Eine lokal installierte eXist-db 5.2.0 Instanz (<a href="http://exist-db.org/exist/apps/homepage/index.html">http://exist-db.org/exist/apps/homepage/index.html</a>)</p>

1 3 : 0 0 - 1 4 : 00	Programmpause
1 4 : 0 0 - 1 4 : 30	<p>Vortrag</p> <p><b>Clemens Neudecker (SBB Berlin): Volltexte und Formate: ALTO und Co. (Folien)</b></p> <p>Der Vortrag stellt die Möglichkeiten und Anforderungen für die Kodierung von mit OCR erzeugten Volltexten am Beispiel des ALTO XML Standard dar. Dabei wird auch auf alternative Formate wie PAGE-XML, hOCR und TEI sowie die Besonderheiten von METS/ALTO für digitalisierte Zeitungen eingegangen.</p> <p><a href="#">Folien</a></p>
1 4 : 3 0 - 1 5 : 00	<p>Vortrag</p> <p><b>Max-Ferdinand Zeterberg (SUB Göttingen): Dateneingabeformulare für RDF Daten auf Basis von SHACL/DASH (Folien)</b></p> <p>Mit Data Shapes [1] gibt es eine SHACL-Erweiterung, die der Erzeugung von Dateneingabeformularen dient. Damit lassen sich die Definition von Formularfeldern einfach mit der Validierung der erzeugten Daten in einem Shapes-Graphen verbinden. In diesem Vortrag wird anhand eines Anwendungsbeispiels gezeigt, wie SHACL/DASH in der Praxis eingesetzt werden kann und welche Herausforderungen dabei zu meistern sind.</p> <p>[1] <a href="http://datashapes.org/dash.html">http://datashapes.org/dash.html</a></p> <p><a href="#">Folien</a></p>
1 5 : 0 0 - 1 5 : 30	Programmpause
1 5 : 3 0 - 1 7 : 00	<p>Interaktive Session</p> <p><b>Standardisierte Rechteinformationen in Metadaten - Grundlagen und Austausch</b></p> <p><b>Leitung:</b> André Hohmann (SLUB Dresden)</p> <p><b>Wiki-Seiten des Vortrags:</b> <a href="#">2021-04-27 DINI AG KIM Workshop - "Standardisierte Rechteinformationen in Metadaten - Grundlagen und Austausch"</a></p> <p>In der interaktiven Session werden die unterschiedlichen Aspekte der Rechteinformationen, Metadatenstandards und Anwendungsbereiche sowie deren Beziehungen zueinander vorgestellt – allerdings mit Schwerpunkt auf Eigenproduktionen der Gedächtnisinstitution. Lizenzierte Ressourcen werden nur teilweise berücksichtigt. Auf dieser Basis schließen sich ein Austausch und Diskussionen an.</p> <p>Auslöser für die Session ist die Erkenntnis, dass in den vielfältigen Bereichen des Informations- und Metadatenmanagements jeweils <a href="#">unterschiedliche Kenntnisse, Vorstellungen und Anforderungen</a> bestehen. Obwohl in den letzten Jahren in unterschiedlichen Anwendungsbereichen Möglichkeiten zur Erfassung oder Verarbeitung der Rechteinformationen eingeführt wurden (DDB, K10plus, ZDB, ...), fehlen für einige Rechteinformationen weiterhin Standards oder Empfehlungen für ein einheitliches Vorgehen. Eine unmittelbare Interoperabilität zwischen allen Bereichen oder Systemen ist ebenfalls noch immer nicht überall möglich.</p> <p>Das Hauptziel der Session ist die Vermittlung eines <a href="#">Überblicks</a> über das Thema Rechteinformation – auf Basis der Erkenntnisse der <a href="#">Lizenzen Gruppe</a>. Weitere Ziele der Session sind (1) die Stärkung des Bewusstseins für standardisierte Rechteinformationen in den unterschiedlichen Anwendungsbereichen, (2) das Werben für die Erfassung von Rechteinformationen in den Metadaten und (3) die Benennung sowie die perspektivische Zusammenführung der Anforderungen in den verschiedenen Anwendungsbereichen. Zudem sollen Impulse für die Gründung von Arbeitsgruppen für Rechteinformationen in den jeweiligen Anwendungsbereichen gegeben werden. Alternativ steht auch die Mitwirkung in der Lizenzen Gruppe zur Verfügung.</p> <p>Vorkenntnisse oder Voraussetzungen sind für die Teilnahme nicht erforderlich.</p>

Mittwoch, 28. April 2021

09 : 30 - 11 : 30	<p>Hands-On Tutorials</p> <p><b>Grundlagen von Datenformaten</b></p> <p><b>Leitung:</b> Jakob Voss (VZG Göttingen)</p> <p>Der Workshop Grundlagen von Datenformaten soll einen praktischen Überblick wesentlicher Gestaltungsmittel von Datenformaten liefern. Anhand von Beispielen und kleinen Übungen werden die Grundbegriffe Syntax, Schema, Standard und Semantik verdeutlicht dazu passende Werkzeuge vorgestellt.</p> <p><b>Vorkenntnisse:</b> Es werden keine Vorkenntnisse vorausgesetzt.</p> <p><a href="#">Folien</a></p>
11 : 30 - 12 : 00	Programmpause
12 : 00 - 13 : 00	<b>Öffentliche Sitzung der DINI AG KIM</b>
13 : 00 - 14 : 00	Programmpause
14 : 00 - 14 : 30	<p>Vortrag</p> <p><b>Steffen Rörtgen (GWDG): Vorstellung der neuen KIM-Gruppe <a href="#">Curricula</a> (Folien)</b></p> <p>Die Curricula-Gruppe der DINI-AG-KIM befasst sich mit der Beschreibung von Curricula im Bildungsbereich und entsprechenden Metadatenstandards.</p> <p>Sie bietet eine Austauschplattform für interessierte Personen sowie Fachleute, die gemeinsam Standards und Beispiele sammeln wollen, um Curricula maschinenlesbar darzustellen.</p> <p>Die Curricula-Gruppe hat sich aus der <a href="#">OER-Metadatengruppe</a> heraus gegründet. Beide Themen sind eng miteinander verwandt, da sie sich mit Metadaten im Bildungsbereich befassen, haben jedoch unterschiedliche Schwerpunkte. Um ihrer jeweiligen Komplexität dauerhaft gerecht zu werden, wurde sich dafür entschieden unter dem Dach der DINI-AG-KIM eine eigenständige Gruppe zu gründen, die sich hiermit befasst.</p> <p>In dem Vortrag möchte ich kurz die Gruppe und in den ersten Sitzungen identifizierte Themen und Arbeitsschwerpunkte vorstellen.</p> <p><a href="#">Folien</a></p>
14 : 30 - 15 : 00	<p><b>Lightning Talks</b></p> <ul style="list-style-type: none"> <li>• Kathrin Schnieders (UOS), Sandra Mierz (TIB): TAPIR - ready to ROR! Nachnutzung von Metadaten aus Datacite Commons (<a href="#">Folien</a>)</li> <li>• Julia Beck (FID DK): "Mediatheken der Darstellenden Kunst digital vernetzen" - Austausch zu audiovisuellen Metadaten (<a href="#">Folien</a>)</li> <li>• Wolfgang Fahl: PIDs und Personenbezogene Daten im Event-Kontext (<a href="#">Kontakt</a>)</li> </ul>
15 : 00 - 15 : 15	<p>Abschluss</p> <p><b>KIM Workshop 2021 - Wrap-Up</b></p> <p>Alex Jahnke (SUB Göttingen), Jana Hentschke (ZBW Hamburg/Kiel)</p>
15 : 15 - 15 : 30	Programmpause

15	Interaktive Session
:	
30	<b>4. Treffen der Selbsthilfegruppe Verlagsmetadaten</b>
-	
17	Im Rahmenprogramm des letztjährigen KIM-Workshop online hat sich die offene Erfahrungsaustauschrunde <a href="#">Selbsthilfegruppe Verlagsmetadaten</a>
:	gegründet. Seitdem finden vierteljährlich 90-minütige Videokonferenzen statt. Alle Interessierten sind herzlich dazu eingeladen, egal ob
00	Wiedersehen oder Kennenlernen!