ZBW Publikationsarchiv

Publikationen von Beschäftigten der ZBW – Leibniz-Informationszentrum Wirtschaft
Publications by ZBW – Leibniz Information Centre for Economics staff members

Kasprzik, Anna; Pohl, Adrian

Contribution to Periodical — Published Version

Tagungsbericht: SWIB23 – 15th Semantic Web in Libraries Conference. Staffelstabübergabe bei den Chairs nach 15 Jahren

b.i.t.online

Suggested Citation: Kasprzik, Anna; Pohl, Adrian (2023) : Tagungsbericht: SWIB23 – 15th Semantic Web in Libraries Conference. Staffelstabübergabe bei den Chairs nach 15 Jahren, b.i.t.online, ISSN 1435-7607, b.i.t.verlag, Wiesbaden, Vol. 26, Iss. 6, pp. 588-591

This Version is available at:
http://hdl.handle.net/11108/604

Kontakt/Contact
ZBW – Leibniz-Informationszentrum Wirtschaft/Leibniz Information Centre for Economics
Düsternbrooker Weg 120
24105 Kiel (Germany)
E-Mail: info@zbw.eu
https://www.zbw.eu/de/ueber-uns/profil-der-zbw/veroeffentlichungen-zbw

Standard-Nutzungsbedingungen:
Dieses Dokument darf zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden. Sie dürfen dieses Dokument nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Sofern für das Dokument eine Open-Content-Lizenz verwendet wurde, so gelten abweichend von diesen Nutzungsbedingungen die in der Lizenz gewährten Nutzungsrechte.

Terms of use:
This document may be saved and copied for your personal and scholarly purposes. You are not to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public. If the document is made available under a Creative Commons Licence you may exercise further usage rights as specified in the licence.

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Mitglied der Leibniz-Gemeinschaft

Tagungsbericht:

SWIB23 – 15th Semantic Web in Libraries Conference

Staffelstabübergabe bei den Chairs nach 15 Jahren

Anna Kasprzik und Adrian Pohl

Die 15. Ausgabe der SWIB (Semantic Web in Libraries) Konferenz fand dieses Jahr erstmals in Berlin und erstmals wieder als Präsenz-Konferenz statt, nachdem sie drei Jahre lang aufgrund der COVID-19-Pandemie online abgehalten wurde. Organisiert von der ZBW – Leibniz-Informationszentrum Wirtschaft und dem Hochschulbibliothekszentrum des Landes Nordrhein-Westfalen (hbz), war die Konferenz in Kooperation mit der Staatsbibliothek zu Berlin – Preußischer Kulturbesitz ein voller Erfolg. 125 Teilnehmende aus 18 Ländern kamen vom 11. bis 13. September 2023 ins spätsommerliche Berlin.

Die Konferenz Semantic Web in Libraries zielt darauf ab, aktuelle Themen und Projekte im Bereich Linked Open Data / Semantic Web und in den letzten Jahren auch Machine Learning für Bibliotheken zu diskutieren. IT-Personal, Entwickler:innen, Bibliothekar:innen und Forscher:innen aus der ganzen Welt kamen zusammen, um ihr Wissen und ihre Erfahrungen auszutauschen.

Sarven Capadisli, Keynote Speaker und Anna Kasprzik

Der erste Tag war als Pre-Conference-Day den Workshops gewidmet. Dazu zählten Workshops, die in das automatische Indexierungstool Annif einführten, der Workshop „Network of Terms", in dem die Teilnehmenden lernten, wie sie ihre Daten durch Verlinkung aufwerten können, oder ein Workshop, in welchem die Teilnehmenden erfuhren, wie sie optimal verwendbares Linked Open Data mit Metafacture erstellen können. Weitere Workshops widmeten sich den Themen "Truly shared cataloguing ecosystem development workshop" und "Nightwatch – metadata management at your fingertips".

Das dezentrale Web als Chance für Bibliotheken?

Dr. Anna Kasprzik (ZBW – Leibniz-Informationszentrum Wirtschaft) eröffnete als Co-Chair des Programmkomitees den zweiten Tag der internationalen Konferenz. Anna Kasprzik löste Joachim Neubert ab, der seit 2008 einer der beiden Chairs der SWIB war.

Die Keynote der SWIB23 hielt Sarven Capadisli. Der Vortrag widmete sich den Herausforderungen im Web und den Möglichkeiten, die das dezentralisierte Web zu deren Überwindung bietet, sowie der Frage, wie Bibliotheken eine Rolle in diesem dynamischen Raum spielen können. Richard Wallis aus Großbritannien sprach über die Erfahrungen des National Library Board of Singapore bei der Implementierung eines operativen Linked-Data-Managementsystems.

Nach den beiden Eröffnungsvorträgen moderierte Joachim Neubert (ZBW) die Session unter der Überschrift „Authorities". Steven Folsom von der Cornell University Library (USA) sprach über die Aufrechterhaltung robuster Lookup-Dienste. Enno Meijers von der National Library of the Netherlands präsentierte das Projekt „Network of Terms", das bereits weitgehend im Bereich des kulturellen Erbes in den Niederlanden angenommen wurde. Michael Lindsey von der Berkeley Law Library hingegen sprach über die Verwendung von Wikibase als Normdatenbank für das institutionelle Repositorium seiner Institution.

Interaktive Lightning Talks und Breakout Sessions

Der zweite Tag der Konferenz beinhaltete eine spezielle Session mit Lightning Talks und interaktiven Breakout Sessions. Diese waren darauf ausgelegt, die Teilnehmenden in fokussierten Diskussionen über aktuelle Fragen und aufkommende Trends im Bereich einzubinden. Hier sind die Highlights jeder Session:

1. Die Zukunft der SWIB gestalten: Die beiden Chairs Adrian Pohl und Anna Kasprzik moderierten diese Session. Es ging darum, Ideen für eine zukünftige internationale und interaktive SWIB zu sammeln, die gleichzeitig inklusiv und klimafreundlich ist. Während es hierfür be-

reits aus der Wikimedia-/Wikidata-Community Vorbilder für dezentrale Modelle des Austausches gibt, kam aus der SWIB-Community doch auch die Rückmeldung, dass persönliche Begegnungen von besonderem Wert seien.

2. Standardisierung von Änderungen an Entitätsdatensätzen: Simeon Warner und Steven Folsom aus den USA gaben eine Einführung in die vorläufige Spezifikation der Entity Metadata Management API (EMM API) durch und sammelten erste Reaktionen und Anwendungsfälle für mögliche zukünftige Versionen der Schnittstelle.

3. Bibliotheken, Wikidata und Linked-Data-Projekte: Will Kent und Eduards Skvireckis führten eine Session durch, um zu erkunden, wie Bibliotheksprofis mit Wikidata arbeiten. Egal, ob es darum geht, die Daten von Wikidata zu nutzen, zu Wikidata beizutragen oder in einer Bibliotheksumgebung Projekte mit Wikidata zu realisieren, die Session lud alle ein, ihre Erfahrungen zu teilen.

4. Best Practices für das Teilen und Entdecken von ETL-Workflows: Tobias Bülte diskutierte Best Practices zur Förderung des Teilens und der Wiederverwendung von ETL-Workflows (Extract, Transform, Load) und stellte Fragen zur Verbesserung der Auffindbarkeit und Nutzbarkeit bestehender Workflows.

5. SoVisu+: Ausgangspunkte und Grundlagen eines nationalen CRIS: David Reymond, Joachim Dornbush, Alessandro Buccheri und Raphaëlle Lapôtre aus Paris (Frankreich) diskutierten das Projekt SoVisu+, ein modulares Open-Source-Current-Research-Information-System (CRIS), das speziell für die Bedürfnisse der französischen Hochschul- und Forschungsgemeinschaft entwickelt wird. Sie präsentierten die Funktionen von SoVisu und die geplante Architektur für SoVisu+ und sammelten Feedback zu Herausforderungen und Möglichkeiten.

Diese interaktiven Sessions ermöglichten es den Teilnehmenden, sich niedrigschwellig mit den Themen auseinanderzusetzen und konkrete Anliegen und Ideen in einer offenen Atmosphäre zu diskutieren.

Datenmodellierung in der Welt der Informationsinfrastruktureinrichtungen

Nach den Lightning Talks ging es am selben Tag weiter mit Vorträgen und Präsentationen. Moderiert von Osma Suominen (National Library of Finland), boten die letzten beiden Vorträge des Tages eine vertiefte Perspektive auf das übergeordnete Thema der Datenmodellierung.

1. Hollinger's Box: Das Abrufobjekt am Rande der Ontologie

Ruth K. Tillman (Penn State University Libraries, USA) und Regine Heberlein von der Princeton University Library (USA) betrachteten die Herausforderungen und Möglichkeiten von Linked-Data-Ontologien im Bereich der Archivbeschreibung. Sie konzentrierten sich auf den „Kontaktbereich" zwischen der intellektuellen und physischen Beschreibung archivierter Ressourcen. Dabei ging es vor allem um den Punkt, an dem das Suchsystem die Verantwortung an ein verwaltendes System überträgt. Sie untersuchten, inwiefern aktuelle und aufkommende Linked-Data-Standards wie „Records in Contexts", die BIBFRAME-Erweiterung für Kunst und seltene Materialien und Linked.Art den Abrufprozess unterstützen können. Dabei stellten sie die Frage, inwieweit diese Standards eine nahtlose Integration mit externen Modellen ermöglichen, die solche Abrufobjekte definieren.

Angeregte Diskussionen in den Workshops

2. Entwicklung der Share-VDE-Ontologie: Ziele, Grundsätze und Prozess

Tiziana Possemato, Jim Hahn und Oddrun Ohren präsentierten die Share-VDE (SVDE)-Initiative, die Bibliothekskataloge und Normdateien in einer gemeinsamen Entdeckungsumgebung zusammenführt. Diese basiert auf der Linked-Data-Ontologie BIBFRAME und wird durch die SVDE-Ontologie erweitert. Der Vortrag beleuchtete den Designprozess und die Ziele der SVDE-Ontologie, die darauf ausgerichtet sind, Konsistenz und Klarheit in der Entdeckungsschicht von SVDE zu bieten. Ein Hauptziel war die Wiederverwendung bestehender Vokabulare, um die Komplexität der SVDE-Ontologie zu reduzieren. Die Ontologie bietet Interoperabilität unter anderem durch direkte Referenzen zwischen BIBFRAME, LRM und RDA.

Beide Präsentationen trugen wesentlich zur Diskussion über die Rolle und die Möglichkeiten der Datenmodellierung in Informationsinfrastruktureinrichtungen bei und boten zahlreiche Anknüpfungspunkte für die Weiterentwicklung und Anpassung bestehender Standards und Praktiken.

Anwendungen von Wikimedia-Plattformen

Der dritte Tag der Konferenz Semantic Web in Libraries konzentrierte sich auf das Thema „Utilizing Wikimedia". Die Vorträge beleuchteten vielfältige Anwendungen und

Raum zum Netzwerken ist ein zentrales Element der SWIB

Strategien, um Wikimedia-Plattformen für Informationsinfrastrukturen zu nutzen.

1. Von EAD zu MARC zu Wikidata und zurück: Werkzeuge und Workflows für Archivmetadaten an den University of Washington Libraries

Crystal Yragui und Adam Schiff präsentierten die Arbeit des Labor Archives of Washington (LAW), einer Kooperation zwischen den University of Washington Libraries und dem Harry Bridges Center for Labor Studies. Mit über 200 Sammlungen ist das Archiv eine der größten Ressourcen für Arbeitsmaterialien in den USA. Seit 2020 wird daran gearbeitet, traditionelle bibliothekarische Metadaten in Wikidata zu integrieren. Ziel ist es, die Forschung durch die Bereitstellung von Linked Open Data zu fördern. Yragui und Schiff erläuterten semi-automatisierte Workflows zur Erstellung von Wikidata-Elementen und diskutierten die Implementierung neuer Wikidata-Eigenschaften für Archivmetadaten.

2. Entwicklung eines Linked-Data-Workflows mit Wikidata

Will Kent von Wiki Education sprach über die zunehmende Bedeutung von Wikidata als Plattform für Linked Data. Er stellte fest, dass es zwar keine standardisierten Anleitungen für die Arbeit mit Wikidata gibt, aber gerade diese Offenheit ermögliche eine Vielzahl von Projektansätzen. Kent teilte Erkenntnisse aus verschiedenen Wikidata-Projekten und gab allgemeine Empfehlungen für die Realisierung von Wikidata-Projekten, insbesondere in Bezug auf Linked Data und die Förderung der Kollaboration zwischen verschiedenen Sammlungen.

3. Verlinkung historischer Dokumente durch OCR mithilfe von Wikidata und Wikipedia

Kai Labusch und Clemens Neudecker von der Staatsbibliothek zu Berlin – Preußischer Kulturbesitz präsentierten ihre Arbeit zur Verlinkung von Namen, Orten und Organisationen in digitalisierten Textsammlungen. Sie nutzten Wikidata als umfassende Wissensbasis und kombinierten diese mit einem Wikipedia-Korpus zur Erstellung eines neuronalen Netzwerks zur Identifizierung von Entitäten in historischen deutschen Texten. Ihre Arbeit demonstrierte die Machbarkeit dieses Ansatzes auf Millionen von Seiten historischer Texte und zeigte, wie die Ergebnisse für die Textklassifizierung verwendet werden können.

Die Session gab einen tiefgreifenden Einblick in die vielfältigen Möglichkeiten der Nutzung von Wikimedia-Plattformen für Informationsinfrastrukturen, insbesondere im Kontext von Linked Data und digitaler Sammlungsverwaltung.

Sammlung und Indexierung

Moderiert von Lars Svensson fokussierte die zweite Session des dritten Tages das Thema „Collections". Diese Session beschäftigte sich intensiv mit verschiedenen Aspekten der Sammlung und Indexierung von Materialien in wissenschaftlichen und kulturellen Einrichtungen.

1. Maschinelle Themenindexierung und darüber hinaus für wissenschaftliche Literatur in der Psychologie bei ZPID

Florian A. Grässle und Tina Trillitzsch vom ZPID (Leibniz-Institut für Psychologie) diskutierten ihre Arbeit an PSYNDEX, einer Referenzdatenbank für psychologische Literatur aus dem deutschsprachigen Raum. Die Vortragenden präsentierten den Übergang von einer lexikalischen Systematik zur maschinenbasierten Software Annif. Sie sprachen über ihre Strategien zur Verbesserung der maschinenbasierten Keyword-Vorschläge, indem sie beispielsweise übermäßig allgemeine Konzepte ausschlossen oder neue Konzepte hinzufügten, die im Trainingsset noch nicht erschienen waren.

2. Implementierung des Albrecht-Haupt-Sammlungsportals basierend auf der allgemeinen semantischen Webanwendung Vitro

Georgy Litvinov, Birte Rubach und Tatiana Walther von der TIB – Leibniz-Informationszentrum Technik und Naturwissenschaften beschrieben die Implementierung und Herausforderungen des Albrecht-Haupt-Sammlungsportals. Die Sammlung besteht aus europäischen Zeichnungen und Drucken und wurde unter Berücksichtigung der speziellen Bedürfnisse von Kunsthistoriker:innen und den Prinzipien des Linked Data indexiert. Eine spezielle Ontologie für bildende Kunst, die GESAH Graphic Arts Ontology, wurde entwickelt und in den Indexierungsprozess integriert.

3. Die Welt der German Digital Library (DDB) und die Grenzen Künstlicher Intelligenz

Mary Ann Tan diskutierte am Beispiel der DDB Ergebnisse der Metadatenanalyse im Bibliotheksbereich mit Fokus auf historischen Hintergründen der Katalogisierung während der Französischen Revolution. Sie betonte die Bedeutung einer automatisierten Bewertung der Metadatenqualität, unter Verwendung von Semantik-Web-

Technologien wie Linked Open Data, im Kontext der Verbesserung der DDB-Suchbarkeit. Tan zeigte die Vorteile der FRBR-Strukturierung des bibliografischen Datensatzes der DDB. Abschließend diskutierte sie Herausforderungen und Lösungen unter Berücksichtigung neuer Entwicklungen in den Bereichen natürliche Sprachverarbeitung, maschinelles Lernen und Wissensgraph-Embeddings.

Insgesamt bot die Session tiefgehende Einblicke in innovative Ansätze und Technologien zur Verbesserung der Sammlung, Indexierung und Verfügbarkeit von Daten in verschiedenen wissenschaftlichen und kulturellen Kontexten.

Die letzte Session der SWIB23 beschäftigte sich mit dem Thema „Aggregators" und wurde von Julia Beck moderiert. In dieser Session wurden zwei interessante Vorträge präsentiert:

1. FranceArchives: Ein Portal für französische Archive

Fabien Amarger von Logilab berichtete über FranceArchives, ein Online-Aggregator-Portal, das den Zugriff auf Metadaten von etwa 140 französischen Archivinstitutionen ermöglicht. Forscher:innen und Enthusiast:innen können hier nach über 21 Millionen Dokumenten suchen. Die in den Archiven erfassten Behörden werden mit Linked-Open-Data-Repositories abgeglichen, um die Daten zu bereichern. Die Daten werden auch im RDF-Format veröffentlicht.

2. Richtung einer Methodik zur Validierung von Metadatenanreicherungen in Europeana

Antoine Isaac von der Europeana Foundation berichtete über die Validierung von Metadatenanreicherungen in der Europeana und ihre entscheidende Rolle. Er stellte eine entwickelte Methodik vor, um die Qualität von Anreicherungen zu bewerten, und beschrieb deren Anwendung bei Geo-Anreicherungen des Projekts Jewish History Tours.

Die SWIB23 war ein inspirierendes Event, das viele spannende Themen und Projekte im Bereich Linked Open Data, Semantic Web und Bibliotheken abdeckte. Neben den technischen Aspekten bot die Konferenz auch eine Plattform für den Austausch von Ideen und Erfahrungen, die für die Bibliothekscommunity von großem Wert sind. Die Themen reichten von dezentralisierten Webansätzen bis hin zu praktischen Anwendungen in Bibliotheken, und es war ein wichtiger Schritt in Richtung einer stärker integrierten und vernetzten bibliothekarischen Zukunft.

Der Vater der SWIB wird verabschiedet

Adrian Pohl (hbz), der zweite Co-Chair des Programmkomitees, verabschiedete nicht nur das Publikum, bedeutete das Ende der SWIB23 doch gleichzeitig das Ende einer SWIB-Ära: Joachim Neubert (ZBW), der 2009 gemeinsam mit Anette Seiler (hbz) die erste SWIB-Konferenz organisiert hatte und seitdem in die Programmgestaltung und Organisation jeder SWIB involviert war, geht Anfang 2024 in den Ruhestand. Joachim Neubert hat somit maßgeblich zum Erfolg der SWIB beigetragen und sie zu dem Forum für Vernetzung und Community-Aufbau und zum Austausch innovativer Ideen und Forschung gemacht, das sie heute ist. Seine Hingabe und sein unermüdlicher Einsatz haben die SWIB zu einer wichtigen Plattform für Fachleute und Interessierte aus der ganzen Welt gemacht. Das Publikum der SWIB23 verabschiedete ihn mit gebührendem Applaus.

Reger Austausch in den Kaffeepausen

Infokasten

- URL: https://swib.org/
- Save the date: Die kommende SWIB24 wird vom 20. bis 22. November 2024 stattfinden.
- Folien und Tagungsmitschnitte sind online verfügbar: https://swib.org/swib23/programme.html

Dr. Anna Kasprzik

Nach einem Studium der Formalen Sprachwissenschaft, Informatik und Kognitionspsychologie und einer Promotion in Theoretischer Informatik absolvierte Dr. Anna Kasprzik das Bibliotheksreferendariat für den höheren Dienst in Konstanz und München. Anna Kasprzik war in einem IT-Projekt für den Bibliotheksverbund Bayern und in der Forschung und Entwicklung an der TIB Hannover tätig und leitet seit 2018 die Automatisierung der Inhaltserschließung mit Machine-Learning-Methoden an der ZBW – Leibniz-Informationszentrum für Wirtschaft.
a.kasprzik@zbw-online.eu

Adrian Pohl

Adrian Pohl arbeitet seit 2008 im Hochschulbibliothekszentrum des Landes Nordrhein-Westfalen (hbz). Er ist Koordinator der Arbeitsgruppe der Open Knowledge Foundation zu Open Bibliographic Data und Vorstandsmitglied der OKF DE
pohl@hbz-nrw.de