

Christine Frodl, Alexander Haffner

Dublin Core 2009

International Conference on Dublin Core and Metadata Applications - »Semantic Interoperability of Linked Data« vom 12. bis 16. Oktober 2009, Seoul, Korea

Die DC-2009 als neunte »International Conference on Dublin Core and Metadata Applications« fand im Oktober 2009 erneut in Asien statt. Nachdem im Jahr 2007 Singapur als Veranstaltungsort gewählt worden war, luden im vergangenen Jahr die National Library of Korea und die Korean Library Association ein.

Der Schwerpunkt »Semantic Interoperability of Linked Data« trifft den Zeitgeist genau und hat neben hoch interessanten Vorträgen auch intensive Diskussionen hervorgebracht.

Die Konferenz wurde u. a. durch Basis-Tutorials begleitet, die durch die Dublin Core Metadata Initiative (DCMI) durchgeführt wurden.

Thomas Baker, DCMI, Chair, referierte über die Entwicklung von Dublin Core, über aktuelle Standards und Modelle sowie Aspekte der Interoperabilität im »complex web of data«. Dabei betonte er den Nutzen und den Einsatz von Resource Description Framework (RDF) in Metadatenstandards für den Weg in das Semantic Web. Des Weiteren hob er die Schnelligkeit von Technologien hervor (»All technology is transitional«) und die Wichtigkeit, sich auf neue Technologien einzulassen, um aktuellen Ansprüchen gerecht zu werden und Daten von einer Generation zur nächsten übernehmen zu können. Als aktueller Anspruch kann insbesondere die Orientierung an »Linked Data« und den damit verbundenen »Data Clouds« im Semantic Web gesehen werden. Verlinkte Daten erlauben es, physisch vorhandene, wie auch abstrakte Dinge, eindeutig mittels URIs zu identifizieren und miteinander in Beziehung zu setzen. Als abstrakter Gegenstand kann z. B. die Beschreibung einer Person angesehen werden, die mit einer digitalen Ressource assoziiert und in die Rolle des Autors gesetzt wird. Für die Ansetzung dieser Infor-

mationen müssen Standards vorhanden sein, um die Kompatibilität zu anderen »Data Clouds« sicher zu stellen und keine isolierte Wolke im Semantic Web zu bilden. Hierfür ist es notwendig, Ontologien durch die Verwendung von RDF-Schema (RDFS) und Web Ontology Language (OWL) und kontrollierte Vokabulare durch Simple Knowledge Organization Framework (SKOS) zu etablieren. Die Berücksichtigung dieser Standards stellt die Grundvoraussetzung für eine Internationalisierung der Bibliothekswelt und für die Nachhaltigkeit der angesetzten Daten im Semantic Web dar. Außerdem berichtete Thomas Baker, dass sich nach Jahren der Zurückhaltung »Mikroformate« und »eingebettete Metadaten« breiterer Anwendung erfreuen. Insbesondere Betreiber von Suchmaschinen haben deren Tragweite erkannt, was eine stärkere Unterstützung und Interpretation dieser Daten nach sich zieht.

Auch der Benutzer rückt ins Zentrum aller Entwicklungen. Diese Aussage wurde nicht nur im Tutorial unterstrichen, sondern zog sich thematisch durch die ganze Konferenz.

In den nachfolgenden Tutorials referierte Marcia Lei Zeng, School of Library and Information Science, Kent State University, Ohio, über nicht-DCMI-Metadaten und deren Interoperabilität. Die Vielfalt an Communities zieht eine große Breite an Anforderungen nach sich. Folglich haben sich verschiedenste Metadatenstandards in der Praxis durchgesetzt. Diese Metadaten zusammenzuführen stellt derzeit, wie wahrscheinlich auch in der Zukunft, eine Herausforderung dar. Auch in diesem Kontext ist der Einsatz von Werkzeugen für die Generierung sowie die Nachnutzung der Daten von höchster Priorität.

In seiner Keynote hob Michael Crandall, University of Washington Information School, den zu beschreitenden Weg ins Semantic Web sowie die Notwendigkeit von Interoperabilität zwischen Communities unterschiedlichster Domänen hervor und ordnete die Aktivitäten der DCMI in diesen Kontext ein. Er verdeutlichte die Schwierigkeit von

Basis-Tutorials

Benutzer steht im Zentrum der Entwicklungen

Keynotes

Daten-Mappings am Beispiel des Informationsverlustes bei der Übersetzung eines Gedichtes. Eun Chul Lee, President, Korea Library Association, stellte in seiner Keynote »Linked data and users in library« die Frage, ob Bibliotheken effizient genug mit ihren Nutzern kommunizieren, rückte damit die Bedürfnisse der Benutzer an die erste Stelle und verdeutlichte die Rolle der Technologien als Werkzeug zur Erfüllung dieser Bedürfnisse. Außerdem verwies er auf die nächste Generation von Benutzern, die nach seiner Einschätzung davon ausgeht, dass Informationen nicht existieren, wenn diese nicht im Web zu finden sind. Die abschließende Keynote von Alistair Miles, CCLRC Rutherford Appleton Laboratory, mit dem Titel »Towards Semantic Web Deployment: Experiences with Knowledge Organisation Systems, Library Catalogues and Fruit Flies« verdeutlichte, welche Möglichkeiten das Semantic Web für verschiedenste Domänen bereithält.

Die Paper Sessions starteten mit einem sehr praxisnahen Bericht von Michael Panzer, OCLC, und Marcia Lei Zeng zur Modellierung von Klassifikationssystemen und kontrolliertem Vokabular. Michael Panzer berichtete über die Erfahrungen von OCLC bei der Abbildung der Dewey-Dezimalklassifikation (DDC). Der Vortrag von Jian Qin, Xiaozhong Liu, Xia Lin und Miao Chen beschäftigte sich mit Fragen der Suchanfrageverfeinerung, um die Präsentation von Suchergebnissen an die Nutzerbedürfnisse anpassen zu können. Kai Eckert stellte sein Paper zum Thema »A Unified Approach for Representing Metametadata« vor und verdeutlichte die Relevanz von Metametadata im Semantic Web, um die Herkunft und Qualität der Daten sicherstellen zu können. Nachfolgend stellte Wooseob Jeong, School of Information Studies University of Wisconsin – Milwaukee, Untersuchungsergebnisse zur Effektivität von Social Tagging am Beispiel von Youtube vor, Jon Mason diskutierte die Frage, inwieweit DC-Metadaten für die Wissensrepräsentation geeignet sind. Der dritte und letzte Konferenztag wurde durch eine Präsentation von Emma Tonkin, UKOLN, University of Bath, eröffnet, in der sie Methoden für die Architektur von Metadaten und korrespondierenden Benutzerschnittstellen für eine adäquate Suchergebnisexploration erläuterte. Sie stellte in

diesem Zusammenhang die gesammelten Erfahrungen und Vorteile des »Multilayered Paper Prototyping« gegenüber anderen nutzerzentrierten Verfahren zur Evaluation und folglich zur Steigerung der Gebrauchstauglichkeit vor. Ayako Morozumi, Satomi Nomura, Mitsuharu Nagamori und Shigeo Sugimoto stellten ein auf dem FRBR-Modell basierendes Metadata Framework für Mangas vor. Sie verdeutlichten dabei den Bedarf eines Datenmodells, um komplexe Sammlungen bibliografisch zu strukturieren. Es sei insbesondere notwendig, Attribute auf Manifestations- und Exemplarebene anzulegen, durch die digitale Ressourcen, beispielsweise auf mobilen Endgeräten besser nachnutzbar werden. Den Abschluss bildete der Vortrag von Cjien-cheng Liu und Chao-chen Chen über die Archivierung und Organisation digitaler Bilder auf Basis des »Embedded Metadata Framework«. Alle Papers können im Internet nachgelesen werden.¹⁾

In den »Special Sessions on Linked data« wurde das Thema aus unterschiedlichen Perspektiven betrachtet: OCLC stellte die Arbeiten von Bibliotheken im Kontext von sozialen Netzwerken vor. Eva Mendez, University Carlos III of Madrid, präsentierte Art und Gebrauch verschiedener Tagging-Stile durch Nutzer. Der Einsatz von Twitter in unterschiedlichsten Zusammenhängen war ebenso Thema wie das »Science collaboration network SCF«, in dem wissenschaftliche Artikel semantisch annotiert werden. Als Kontrast wurde in der Session »Linking formal data« durch Jane Greenberg, School of Information and Library Science University of North Carolina at Chapel Hill, Manning Hall, über Linking Data im Dryad Project und durch Hans Overbeek, über die Verwendung in der holländischen Regierung berichtet. Antoine Isaac und Makx Dekkers stellten das Datenmodell von Europeana vor. Europeana ist das erste Projekt, welchem es erfolgreich gelang, Metadaten aus Bibliotheken und Archiven zu vereinen und durch Linked Open Data der Öffentlichkeit zugänglich zu machen. Die weiteren Special Sessions richteten den Fokus auf die technischen Aspekte von Linked Data. Ausgehend von den DCMI Terms wurde der Einsatz von Description Set Profiles entsprechend dem DCMI Abstract Model erläutert. Jutta Treviranus, Adaptive Technology Resource Center, University of Toronto, betonte die Wichtigkeit, Men-

Datenmodell zur Strukturierung ist nötig

Special Sessions

Paper Sessions

schen mit besonderen Bedürfnissen, wie z. B. Menschen mit Behinderung, in die Entwicklungen einzubeziehen und ihre Anforderungen zu berücksichtigen. Ebenso wurde durch Karen Rollitt die Notwendigkeit der Internationalisierung zur Minimierung von Barrieren betont. Abschließend wurden Prinzipien für die agile Entwicklung von Application Profiles auf Basis des Singapore Framework vorgestellt. Eine weitere Special Session wurde mit dem Vortrag von Tim Berners-Lee zum Thema Web 3.0 eröffnet. Er rief zur domänenübergreifenden Bereitstellung von »raw data« und der komplexen Vernetzung dieser Daten auf. Er betonte neben den vielfältigen Potenzialen die Notwendigkeit einer zeitnahen Realisierung. Eric Miller präsentierte nachfolgend die vorstellbare Zukunft im Web 3.0. In der anschließenden Diskussion kristallisierte sich die Meinung heraus, dass schnellst möglich verlinkte Daten erstellt werden sollten, die jedoch von hoher Qualität sein müssen. Ergänzend zeigte David Wood technologische Realisierungskonzepte für IT-Landschaften im Semantic Web auf.

Themenkomplex RDA

In der DC Libraries Community Session präsentierte Christine Frodl die aktuellen Arbeiten der DCMI/RDA Task Group und der FRBR Review Group rund um das Thema RDA. Alexander Haffner unterstrich in seinem Vortrag »RDA and the Semantic Web« die Bedeutung von RDA als Katalysator für die Internationalisierung und Interoperabilität von Metadaten. Durch eine konsequente Anwendung von RDA werden Endnutzer in globalisierten Katalogen Suchresultate präsentiert bekommen, die explorativ zu erkunden und an die menschliche Wahrnehmungsfähigkeit angepasst sind. Durch die korrespondierenden RDA-Ontologien und adäquate Wertevokabulare erwächst die

Möglichkeit, die qualitativ hochwertigen Metadaten aus der Bibliothekswelt für die allgemeine Nachnutzung durch Informationssystembetreiber bereitzustellen. Im Optimalfall wird auf Basis von RDA eine »Data Cloud« mit vertrauenswürdigen Daten im Semantic Web entstehen, die als maßgebliche Quelle für die Verlinkung dienen wird. Daher ist es entscheidend, diese Position zeitnah einzunehmen, bevor sie durch konkurrierende kommerzielle Lösungen übernommen wird. Falls die Bibliothekswelt diesen Schritt zu spät geht, wird es nur noch schwer möglich sein, eine entscheidende Rolle im Semantic Web einzunehmen. Außerdem beschäftigte die DC Libraries Community das weitere Vorgehen mit dem DC-Lib Application Profile. Insbesondere wurde eine »FRBRisierung« für das zugrundeliegende Datenmodell intensiv diskutiert. Auch wurde die Frage gestellt, ob ein solches Application Profile in Hinblick auf RDA überhaupt noch notwendig sei. Die Antwort lautete – ja. Grund dafür ist die noch offene Frage nach dem Umgang mit »Altdaten«, die nicht nach FRBR überführt werden. Außerdem wurde die vergleichsmäßige Einfachheit für den Austausch von Metadaten außerhalb des Semantic Webs, wie beispielsweise mittels OAI-PMH hervorgehoben.

Ein wichtiges Fazit am Ende der Konferenz ist die Tatsache, dass die Daten die technologischen Entwicklungen »überleben« werden. Umso wichtiger ist es, qualitativ gute Metadaten zu produzieren und diese mit den dargestellten Methoden und Standards noch »smarter« zu machen.

Die nächste International Conference on Dublin Core and Metadata Applications wird vom 20. bis 22. Oktober 2010 in Pittsburgh, Pennsylvania, USA, stattfinden.

Fazit

Anmerkung

1 <<http://dcpapers.dublincore.org/ojs/pubs/issue/current>>