Zeitschrift: Arbido

Herausgeber: Verein Schweizerischer Archivarinnen und Archivare; Bibliothek Information Schweiz

Band: - (2013)

Heft: 4: Linked Open Data Big Data : alles vernetzt = tout en réseau = tutto in rete

Buchbesprechung: Das Ende des Zufalls : wie Big Data uns und unser Leben vorhersagbar macht [Rudi Klausnitzer]

Autor: Blumer, Eliane

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 07.07.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

Sozialwissenschaften (InFoLiS)». InFoLiS ist ein DFG-gefördertes Projekt des Leibniz-Instituts für Sozialwissenschaften, der Universitätbibliothek Mannheim und des Lehrstuhls für Künstliche Intelligenz an der Universität Mannheim.

Pascal Christoph bringt uns die Datenanreicherung auf LOD-Basis näher. Die Herstellung von Kontext kann als Kataloganreicherung einen Mehrwert bei der Recherche für den Nutzer mit sich bringen. Durch Linked Open Data gehen wir einen Schritt weiter. Es ist nicht nur möglich, Daten zu integrieren, um damit den eigenen Katalog anzureichern. Vielmehr ist die Basiszielsetzung von LOD auch das Publizieren der Daten, wie der Autor zutreffend ausführt. Sein Beitrag beleuchtet einführend die theoretischen Aspekte, die im Zusammenhang mit der Katalog-anreicherung auf Basis von Linked (Open) Data wichtig sind. Anhand des vom Hochschulbibliothekszentrum des Landes Nordrhein-Westfalen (hbz) betriebenen LOD-Services lobid.org wird aufgezeigt, wie bibliografische Daten auf Basis anderer LOD-Quellen angereichert werden können. Abschliessend betrachtet der Verfasser künftige Chancen und Möglichkeiten der Kataloganreicherung. Er erläutert die drei möglichen Ansätze «blosse Verlinkung», «dynamische Kataloganreicherung» und «Datenübernahme mit ihren Vor- und Nachteilen». Schnell zeigt sich, dass die Lizenzierungsfrage bei der Entscheidung, welchen Ansatz die Bibliotheken wählen können, eine zentrale Rolle spielt, da sie bei der Nutzung von Fremddaten von den vertraglichen Nutzungsbedingungen abhängig sind. Ein guter Beitrag, der aber Vorkenntnisse in der Computerlinguistik voraussetzt.

Mit zwei weiteren Beiträgen gehen Markus M. Geipel, Christoph Böhme, Julia Hauser und Alexander Haffner («Herausforderung Wissensvernetzung: Impulsgebende Projekte für ein zukünftiges LOD-Konzept der Deutschen Digitalen Bibliothek») sowie Klaus Ceynowa, Matthias Gross, Andreas Kahl und Gabriele Messmer («Linked Open Data geht in die Fläche: Der B3Kat stellt seine Daten frei») Fragen der Umsetzung an Projekten wie der Deutschen Digitalen Bibliothek und in Länderverbundprojekten nach. Den gängigen bibliografischen Metadatenformaten liegt ein anderes Datenkonzept zugrunde als beim Resource Description Framework (RDF). Einen MARC-, Pica- oder MAB-Datensatz kann man sich als eine Art digitale Karteikarte vorstellen, deren Felder ausgefüllt sind. RDF hingegen betrachtet Daten von bibliografischen Angaben als Netzwerk. Es existieren keine einzelnen Datensätze in herkömmlichem Sinne mehr, sondern nur noch Knoten, die durch Aussagen verknüpft sind. Dieser Paradigmenwechsel macht die Umwandlung von bibliografischen Daten ins RDF schwierig, da noch keine standardisierten Konkordanztabellen zwischen den Bibliotheksformaten und dem RDF geschaffen wurden.

Diese Neuerscheinung sticht in zweierlei Hinsicht aus den bisherigen Publikationen der Reihe «Bibliotheks- und Informationspraxis» des Verlages De Gruyter hervor. Einerseits, da die Autoren von der bisherigen Konzeption des «Ready for use» der Reihe abrücken, was teilweise der Komplexität des Publikationsgegenstandes geschuldet ist, und andererseits, weil es das erste Werk in dieser Reihe des Verlages ist, welches von den Verlegern Open Access unter Creative Common Licence gestellt wird (http://www.degruyter.com/viewbook-toc/product/181080). Das verdient Anerkennung. Ein Gewinn der Reihe «Bibliotheks- und Informationspraxis» ist es bisher gewesen, dass man die Bücher auch als Berufspraktiker ohne allzu spezielle Vorkenntnisse lesen konnte. Daran sollte man seitens des Verlags weiterhin festhalten, denn das ist die eigentliche Stärke der Reihe. Wer sich andererseits tapfer auch ohne linguistische Vorkenntnisse durch die mehr theoretischen Beiträge des Buches durchgebissen hat, legt das Buch mit Gewinn zur Seite, da er die Mög-lichkeiten von (Open) Linked Data erahnt und sich der Schwierigkeiten bei deren Umsetzung bewusst wird.

Big Data für Anfänger – ein Reiseführer für «Nichtinformatiker»

Eliane Blumer

Rudi Klausnitzer: Das Ende des Zufalls – Wie Big Data uns und unser Leben vorhersagbar macht (Ecowin Verlag, Salzburg, 2013)

Das Buch «Big Data. Das Ende des Zufalls» ist der Blick eines «alten Hasen» aus der Kommunikations- und Medienwelt, der es sich zur Aufgabe gemacht hat, tiefer in die Big-Data-Welt einzutauchen und seine Entdeckungsreise mittels eines Buches festzuhalten. Dabei will er seiner Auffassung, dass Big Data eines der Jahrhunderthemen ist, jedoch noch nicht die nötige Aufmerksamkeit erhalten hat, Nachdruck verleihen.

Die Reise von Rudi Klausnitzer beginnt ganz vorne – nämlich bei der Fragestellung, warum denn Big Data? Hier seine Erklärung zusammengefasst: Weil der Mensch, neurologisch bewiesen, den Zufall nicht mag und ihn deshalb so oft als möglich vermeiden will. Und Big Data kann uns dabei helfen, denn durch die maschinelle Kontextualisierung der richtigen Datensets können Vorhersagen getroffen werden, die denn Zufall quasi ausschalten. So kann beispielsweise die Verlinkung von Fahndungsdaten mit bereits existierenden kriminellen Mustern dazu dienen, mögliche Verbrechensorte vorauszusagen und den effizienten Einsatz von Polizisten

bedeuten. Selbstverständlich war dies schon vor der Erfindung von Semantic-Web-Technologien möglich, bedurfte jedoch mehr Zeit und nur menschlicher Denkarbeit. Selbstverständlich wird dies auch nach dem ersten Rush rund um Big und Linked Data möglich sein und bedarf weiterhin menschlicher Denkarbeit, welche einfach durch den Computer unterstützt wird.

Ob «Datenabfall», «Data Scientist», «Big Data», «Dark Data» oder gar «Smart Living», es ist sicher, dass sich Rudi Klausnitzer sehr gut ins Thema der «Grossen Daten» eingearbeitet hat und mit den Begriffen umzugehen weiss. Treffsicher führt er diese im Verlauf der Lektüre ein, erklärt beispielsweise, dass unter «Smart Living» unter anderem intelligente Wohnungen verstanden werden, die den Energieverbrauch messen, und verwendet sie weiter bis zur letzten Zeile. Sein Buch liest sich einfach, gibt einen schnellen Einblick ins Thema der grossen Daten und ist vor allem gespickt mit veranschaulichenden Beispielen, die sowohl die negativen wie auch positiven Auswirkungen oder einfach die Einsatzmöglichkeiten von Big Data aufzeigen. Dabei scheint Klausnitzer Unvoreingenommenheit wichtig zu sein, was durch die jeweils gleiche Anzahl positiver wie negativer Beispiele sichtbar wird. Die Beispiele holt er beispielsweise aus dem Gesundheitswesen mit der Predictive Medicine oder aber dem Marketing, wo Einkäuferdaten analysiert und dementsprechend auf den Einkäufer zugeschnittene Werbeprodukte angepriesen werden.

Der etwas ungeduldige Leser wartet aber bis in die Hälfte des Buches, um eine Definition von Big Data zu erhalten. Einmal bis dorthin gelangt, wird ihm dann aufgezeigt, wo heutzutage überall Daten generiert werden– durch Smartphone, Auto, Haushaltsgeräte oder Applikationen wie «GoogleGlass» – und wie diese durch Algorithmen in Kontext gebracht, zu Smart Data werden können. Schade ist dabei, dass dem Leser bis ans Ende der Algorithmus als nur von «Informatikern», «Mathematikern» und «Statistikern» zu verstehende Rechnungsformel dargestellt wird, die die Intelligenz in die Daten bringt, jedoch nie ein eigentlicher Algorithmus im Buch dargestellt und ausführlicher erklärt wird. Am Ende hat es den Anschein, als ob die Algorithmen die Lösung für alles seien, jedoch nur oben genannte «Auserwählte» ihren wirklichen Hintergrund verstehen können.

Im populärwissenschaftlichen Stil verfasst, soll das Buch ein möglichst breites Publikum ansprechen, denn, wie Klausnitzer selbst im Buch erwähnt: «Bei Big Data geht es um das reale Leben. Da geht es um uns.» Die schnelle Entwicklung der Technologie führt dazu, dass so bald als möglich Fragen zu Datenschutz, Datenerbe oder auch zu Datenkriegen gestellt werden müssen. Klausnitzer betont, dass eine Diskussion bereits stattfindet, dass sich aber jeder darin einbringen sollte, auch um den eigenen Standpunkt zu klären. Dies entspricht der Positionierungsdiskussion der Bibliotheken gegenüber «der Datenfrage» und ist deshalb auch richtig zu erwähnen.

Klausnitzer kommt, wie andere auch, zum Schluss, dass die neuen Trendberufe «Data Scientist», «Data Designer», «Data Journalisten» oder gar «Datatainer» sein werden. Kurz gesagt, Fachkräfte, die mit den Daten umgehen können, sie analysieren, visualisieren, beschreiben beziehungsweise spielerisch weiterverwenden. Dies muss aber nicht heissen, dass nur noch solche Leute ausgebildet werden sollen, sondern zeigt eine neue Entwicklung der Professionen auf.

In einer leicht an Science Fiction anmutenden Zukunftsvision zeigt Klausnitzer auf, dass wir uns auf eine hybride Zeit hinbewegen, in welcher alltägliche Geräte wie Smartphones aber auch der Mensch noch mehr miteinander vernetzt werden. Dabei gibt es drei Lebewesen, «Hybrids», «Algos» und «MTurcs». Erstere sind eine Mischung aus Mensch und Maschine, zweitere sind die Computer und letztere sind diejenigen, die den Maschinen zuarbeiten und von ihnen kontrolliert werden.

Es kann gesagt werden, dass das Buch von Klausnitzer als Reiseführer durch die Welt von Big Data taugt und die aktuellen Beispiele von bereits existierenden Big-Data-Anwendungen aus der Praxis auch im etwas erfahreneren Leser Neugier für mehr wecken. Dennoch mangelt es ihm manchmal an weiterführenden Definitionen oder tiefergehenden, auch vom Autor selbst vertretenen, Ansichten und Auseinandersetzungen mit der kritischen Thematik. Man merkt bei der gesamten Lektüre den Medienprofi, der weiss, wie er Neuigkeiten an den Mann und die Frau bringen muss, sich selbst und seine Meinung dabei so weit als möglich zurückhält und welche Beispiele dabei möglichst viele Menschen ansprechen, wie das Predictive Policing oder vermehrt auf Social Media ausgetragene Wahlkämpfe. Vergessen geht bei der Aneinanderreihung dieser Beispiele leider etwas, die verschiedenen Standpunkte innerhalb einer solchen Bewegung anzusprechen, damit die Beteiligung der «Nichtinformatiker», «Nichtstatistiker» und «Nichtmathematiker», wie es ja vom Autor auch gewünscht ist, wirklich erreicht werden kann. Denn, wie es im Buch zu lesen ist, ist man sich in vielen Bereichen des Potenzials von Big Data bereits bewusst und agiert dementsprechend, es bedarf deshalb sicherlich detaillierterem Hintergrundwissen und der Bereitschaft, um einen Beitrag dazu machen zu können. Abschliessend kann gesagt werden, dass sich das Buch für alle empfiehlt, die schnell einen weiten, aber an der Oberfläche bleibenden Überblick zu Big Data haben möchten. Wer sich, gerade auch an den informationswissenschaftlichen Bereich denkend, vertieft mit den Technologien hinter Big Data oder den eigentlichen Anwendungssprachen auseinandersetzen möchte, für den hat das Buch möglicherweise etwas zu wenig vertiefenden Inhalt.