

Jakob Voß

Was sind eigentlich Daten?

2013

<https://doi.org/10.25969/mediarep/4093>

Veröffentlichungsversion / published version
Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:

Voß, Jakob: Was sind eigentlich Daten?. In: *LIBREAS. Library Ideas* (2013), Nr. 23, S. 4–11. DOI: <https://doi.org/10.25969/mediarep/4093>.

Erstmalig hier erschienen / Initial publication here:

<https://nbn-resolving.org/urn:nbn:de:kobv:11-100212657>

Nutzungsbedingungen:

Dieser Text wird unter einer Creative Commons - Namensnennung 4.0/ Lizenz zur Verfügung gestellt. Nähere Auskünfte zu dieser Lizenz finden Sie hier:

<https://creativecommons.org/licenses/by/4.0/>

Terms of use:

This document is made available under a creative commons - Attribution 4.0/ License. For more information see:

<https://creativecommons.org/licenses/by/4.0/>

WAS SIND EIGENTLICH DATEN?

von Jakob Voß

Daten und ihre Nutzung als Linked Data, Big Data, Forschungsdaten, etc. erfahren aktuell verstärkte Aufmerksamkeit. Bei genauerer Betrachtung lässt sich dieses Interesse bis zu den Anfängen der Computernutzung zurückverfolgen, jedoch ohne dass ein eigenständiges, theoretisches Konzept von 'Daten' erkennbar wäre. Der folgende Beitrag gibt eine kurze Einführung in verschiedene Datenbegriffe aus philosophischer Sicht und verbindet diese mit dem Konzept des Dokuments in der Bibliotheks- und Informationswissenschaft.

Informationen und Daten

Informationen stehen im Mittelpunkt der Forschung verschiedener Disziplinen wie Informatik, Informationswissenschaft und Information Systems. Daten werden dabei höchstens als sekundäres Konzept herangezogen oder quasi synonym zu Informationen aufgefasst.

Bekannt ist vor allem die Informationstheorie von Claude Shannon (1948), der Information mathematisch als eine Menge von Bits definiert und dabei Aspekte der Bedeutung explizit ausklammert. Eine alternative Definition von Information lautet ‚Daten mit Bedeutung‘ (Floridi 2005; Kettinger und Li 2010). Ob Daten Bedeutung haben oder nicht und wie sich damit Daten von Informationen unterscheiden, ist also unklar. Mindestens bis in die 1960er lässt sich zeigen (Gray 2003), dass in der Fachliteratur keine Unterscheidung zwischen Daten und Informationen getroffen wurde. Während später die Bezeichnung ‚Information‘ die Rede von ‚Daten‘ verdrängt, geht man oft noch immer implizit davon aus, dass beide mehr oder weniger gleichzusetzen sind (Boell und Cecez-Kecmanovic 2011; Jr. und II 2010). Während dabei Informationen in verschiedenen Disziplinen eine zentrale Rolle spielen, führend Daten als *primärer* Forschungsgegenstand noch immer ein Schattendasein. Der Ruf nach einer Wissenschaft von Daten, wie vom Informatiker Peter Naur (1966) unter der Bezeichnung „Datalogy“ gefordert, blieb weitgehend ungehört. Erst in den letzten Jahren ist neben rein affirmativen Trends wie Linked Data und Big Data auch eine kritische Auseinandersetzung mit der Frage zu beobachten, was Daten eigentlich sind (siehe beispielsweise der Sammelband von Gitelman (2013)). Zur Klärung des Datenbegriffs kann, wie im Folgenden dargelegt, die Philosophie helfen.

Philosophie der Daten

Wie die lateinisch Wurzel ‚datum‘ (von ‚dare‘ für ‚geben‘) anzeigt, werden Daten oft als ‚etwas Gegebenes‘ angenommen. Der Begriff ‚data‘ geht auf das Werk ‚*Ἀεδομένα*‘ des Euklid zurück. Er zeigt darin, wie gesuchte geometrische Eigenschaften von anderen gegebenen Größen abgeleitet werden können. Mit Beginn des 20. Jahrhunderts wurden im Rahmen erkenntnistheoretischer Überlegungen unmittelbar zugängliche Wahrnehmungen als ‚Sinnesdaten‘ bezeichnet. Eine philosophische Auseinandersetzung mit dem modernen Datenbegriff, wie er im Zusammenhang mit automatischer Datenverarbeitung auftritt, findet sich erst Anfang des 21. Jahrhunderts.

Floridis diaphorische Definition von Daten

Ogleich der Schwerpunkt seiner Analysen auf der Philosophie der Information liegt (Floridi 2010; Floridi 2009), hat der italienische Philosoph Luciano Floridi mit seiner *diaphorischen Definition von Daten* (vom Griechischen ‘διαφορά’ für ‚Unterschied‘) auch Grundlegendes zur Definition von Daten beigetragen. Floridi (2010) definiert Daten als:

x being distinct from *y*, where *x* and *y* are two uninterpreted variables and the relation of ‘being distinct’, as well as the domain, are left open to further interpretation.

Die Definition ist demnach verknüpft mit dem Problemen der Identität und Gleichheit. Wo Dinge praktisch das Gleiche sind oder ununterscheidbar werden, sind also keine Daten möglich. Floridi führt weiter verschiedene Arten von Daten aus, im speziellen

- Unterschiede in der Wirklichkeit (*data de re*). Diese Unterschiede sind nicht direkt zugänglich sondern müssen als gegeben vorausgesetzt werden, da wir ohne sie nichts anderes als Gleichförmigkeit wahrnehmen würden.
- Unterschiede in der Wahrnehmung (*data de signo*) von verschiedenen physikalischen Zuständen.
- Unterschiede zwischen Symbolen (*data de dicto*), beispielsweise zwischen den Buchstaben ‚A‘ und ‚B‘.

Während *data de re* ein rein theoretisches Konstrukt zur Beschreibung der Realität sind, lassen sich *data de signo* und *data de dicto* auch grob als analoge bzw. digitalen Daten interpretieren. Der Großteil von Daten besteht heute aus unterschiedlichen Symbolen die von Computern verarbeitet werden (*data de dicto*). Daten als Unterschiede in der Wahrnehmung physikalischer Zustände spielen im Rahmen der Datenverarbeitung lediglich für Sensoren und Messinstrumente eine Rolle.

Vorstellungen von Daten nach Ballsun-Stanton

Einen weniger theoretischen Ansatz zur Definition von Daten verfolgt der australische Philosoph Brian Ballsun-Stanton. In seiner Dissertation (2012; 2010) arbeitete er mit Mitteln der praktische Philosophie heraus, welche grundlegenden Konzepte unter ‚Daten‘ verstanden werden. Im Wesentlichen lassen sich demnach drei Vorstellungen unterscheiden:

- In der Vorstellung von *Daten als Fakten* (“data as facts”, “data as hard numbers”) sind Daten objektive, reproduzierbare Ergebnisse von Messungen, die wahre Aussagen über die Realität liefern. Beispiele für diese naturwissenschaftliche Sicht auf Daten sind Größenangaben, wie die Masse der Erde oder die Einwohnerzahl einer Stadt.
- In der Vorstellung von *Daten als Beobachtungen* (“data as observations”) sind Daten aufgezeichnete Wahrnehmungen. Sie sind prinzipiell subjektiv, benötigen

Kontextwissen und müssen gefiltert werden, um aus ihnen relevante Informationen zu extrahieren. Beispiele für diese Daten sind die Notizen eines Forschers, Interviews oder Audio- und Videoaufzeichnungen. Ebenso wie in der Vorstellung von Daten als Fakten wird davon ausgegangen, dass sich Daten auf reale Phänomene beziehen; der Bezug ist allerdings nicht direkt sondern muss erst durch Interpretation der Daten wieder hergestellt werden.

- In der Vorstellung von *Daten als binäre Nachrichten* (“data as bits”, “data as communications”) sind Daten Zeichen, die zur Kommunikation dienen. Statt des Bezugs von Daten zur Realität steht in dieser Vorstellung der semiotische Charakter von Daten im Vordergrund. Obgleich Daten als Nachrichten in unterschiedlichen Repräsentationsformen auftreten können (Texte, Diagramme, Tabellen etc.) sind Daten dieser Vorstellung nach letztendlich eine Folge von Bits.¹

Daten als Paradigma

Beim Vergleich mit den aktuellen Diskursen um Linked Data, Big Data und Forschungsdaten lässt sich feststellen, dass bei jedem Trend eine andere Sicht auf Daten dominiert: im *Semantic Web* werden Daten vorrangig als kombinierbare Fakten (oder Faktenaussagen) aufgefasst — symptomatisch für die dabei vorherrschende Vorstellung von Daten ist die Übernahme des Begriffs ‘Ontologie’, der sich auf das Seiende bezieht. Eine Kontextualisierung der Daten, beispielsweise durch Einbindung von Provenienzinformationen, ist innerhalb des RDF-Modells nur sehr schwer möglich.² Im Rahmen von *Big Data* spielt die Interpretation und Filterung von Daten eine größere Rolle. Daten werden hierbei eher als aufgezeichnete Beobachtungen aufgefasst, die sich unter anderem mittels statistischer Verfahren wie Textmining und Bildverarbeitung auswerten lassen. Eine einheitliche Vorstellung von *Forschungsdaten* lässt sich angesichts der Unterschiede zwischen Fachdisziplinen weniger klar festmachen. Geht man davon aus, dass unter Forschungsdaten alle Daten subsumiert werden können, die im Rahmen einer wissenschaftlichen Tätigkeit anfallen (vgl. Walther u. a. 2013), so steht hier die Funktion von Daten innerhalb eines kommunikativen Prozesses im Vordergrund. Forschungsdaten folgen damit der Vorstellung von Daten als binären Nachrichten.

Wissenschaft von Daten

Die semiotische Sicht, nach der sich Daten als digitale Nachrichten auffassen lassen, ermöglicht eine Auseinandersetzung mit Daten, die diese ins Zentrum der Bibliotheks- und Informationswissenschaft rückt. Die fortschreitende Digitalisierung aller

¹ Eine genauere Untersuchung dieser verschiedenen Formen von Daten findet sich in meiner Dissertation (Voß 2013; Voß 2011).

² Übliche Strategien beinhalten eine Erweiterung des Tripel-Modells oder die Kodierung der einzelnen Tripel-Bestandteile durch andere Tripel (Reification).

Kommunikations- und Publikationsmittel hat dazu geführt, dass das, was Bibliothek und Informationseinrichtungen sammeln, ordnen und verfügbar machen, letztlich Daten sind. Die Forderung nach Neuorientierung an einer 'Datenwissenschaft' wäre jedoch irreführend, da dieser statistisch orientierten Fachdisziplin ein anderer Datenbegriff zugrunde liegt: das primäre Ziel der Datenwissenschaft ist die Auswertung von Daten, um daraus neue Informationen zu gewinnen; Daten werden dabei vor allem als Messungen und Beobachtungen aufgefasst, aus denen sich Unterschiede zwischen physikalischen Zuständen oder anderen realen Gegebenheiten ableiten lassen. Für Bibliotheken und Informationseinrichtungen ist dagegen eine grundsätzlich andere Eigenschaft von Daten von Interesse, und zwar ihr Charakter als publizierte Nachricht. Der Fachbegriff für solche Nachrichten ist das 'Dokument' und die dazugehörige Wissenschaft die Dokumentation oder Dokumentationswissenschaft.

Das Dokument im digitalen Zeitalter

Die Beschäftigung mit Dokumenten reicht bis zu den Ursprüngen der Bibliotheks- und Informationswissenschaft zurück. Bereits in den Anfängen der Informationswissenschaft war auch klar, dass Dokumente nicht an bestimmte physische Träger gebunden sind (vgl. Rayward 1994; Otlet 1934). Mit dem Aufkommen von Computern zur Verarbeitung bibliographischer Daten wechselte allerdings die Aufmerksamkeit vom 'Dokument' zur 'Information'. Diese Negation des Dokuments findet sich auch außerhalb der Fachdisziplin und ist tief im Verständnis heutiger Informationssysteme verankert. So schreibt Tim Berners-Lee (1999), der Erfinder des WWW, in seiner Referenz auf das erste Hypertextprojekt Xanadu von Ted Nelson:

Ted described a futuristic project, Xanadu, in which all the world's information could be published in hypertext.

Ted Nelson, dessen Konzept von Hypertext in wesentlichen Punkten über heutige Hypertexte hinausgeht, korrigierte diese Beschreibung in einer Replik (Nelson 2010):

Not 'all the world's information', but all the world's documents.

Zwar sind wir in der Regel vor allem an Informationen interessiert; ohne Dokumente lassen sich Informationen aber weder ausdrücken noch kommunizieren. Aus diesem Grund werden digitale Dokumente und Daten oft fälschlicherweise als (Sammlungen von) Informationen bezeichnet oder mit Informationen gleichgesetzt. Gleichzeitig erodieren im Digitalen die Grenzen des Dokumentes (Dudek 2012), da praktisch jede Teilmenge, Obermenge oder Veränderung eines Dokumentes ebenso als Dokument aufgefasst werden kann.³ Prinzipiell bildet jedes aus Daten bestehende Objekt (Datei, Datensatz, Webseite, etc.) ein Dokument, sobald es als publizier- oder kommunizierbares Objekt aufgefasst wird — eine grundsätzlich Unterscheidung zwischen publizierten oder kommunizierten Daten und digitalen Dokumenten ist deshalb nicht möglich.

³ Mit seinem Hypertextsystem Xanadu versucht Ted Nelson zu zeigen, dass Digitalität nicht nur Dokumente auflöst sondern auch dazu dienen kann, Spuren zwischen Dokumenten besser sichtbar zu machen (@Nelson1981; Nelson 2010).

Während sich einerseits der traditionelle Dokumentbegriff mit der Digitalisierung auflöst, lässt sich andererseits ein verstärktes fachliches Interesse am Dokument feststellen. Angestoßen wurde diese Rückbesinnung auf den Kerngegenstand der ehemaligen Dokumentationswissenschaft unter anderem durch Beiträge von Michael Buckland, mit der Frage “What is a ‘document’” (Buckland 1997) bzw. “What is a ‘digital document’” (Buckland 1998).⁴ Buckland bezieht sich im Wesentlichen auf die Arbeit der französischen Dokumentarin Suzanne Briet (1951), nach der prinzipiell jeder Gegenstand als Dokument erhalten kann, vorausgesetzt er wird “preserved or recorded, intended to represent, to reconstruct, or to demonstrate a physical or conceptual phenomenon.”⁵ Das Dokument wird also nicht über seinen Inhalt sondern über seine Funktion zur “Repräsentation, Konstruktion oder als Nachweis” seines Inhalts definiert.

Die Definition von Dokumenten über ihre Funktion als Zeichen spricht auch für eine Zusammenführung von digitalen Dokumenten mit dem kommunikativen Datenbegriff: werden Daten als binäre Nachrichten aufgefasst, so sind praktisch alle Daten auch Dokumente. In der Terminologie der Datenbeschreibung lässt sich dieser Zusammenhang auch so ausdrücken, dass Daten genau dann Dokumente sind, wenn sie in einem Format mit einem sinnvollen Datenmodell vorliegen (Voß 2009; Voß 2013) und somit bedeutungstragende Zeichen im semiotischen Sinne sind.

Die Vorstellung von Daten als kommunizierten Bits (Balloun-Stanton) oder anderen Symbolen (Floridi) hat zum Vorteil, dass der Datenbegriff ohne Rückgriff auf den Informationsbegriff auskommt und mit dem Begriff des digitalen Dokuments zusammenfällt. Während bei Daten als Fakten und Daten als Beobachtungen die Frage im Vordergrund steht, welche Informationen in den Daten „enthalten“ sind, ist bei Daten als digitalen Dokumenten deutlich, dass je nach Rezeptionskontext ganz unterschiedliche Inhalte im Vordergrund stehen können. So kann der gleiche Datensatz bzw. das gleiche digitale Dokument je nach Kontext unterschiedliche Informationen liefern.

Angesichts dieser Uneindeutigkeit ist es für Bibliotheken und Informationseinrichtungen hilfreich, dass es nicht ihre primäre Aufgabe ist, Dokumente auszuwerten, um daraus neue Informationen zu gewinnen. Die Beschäftigung mit den Inhalten von Dokumenten, ganz gleich ob digital oder physisch, ist nämlich Aufgabe der Autoren und Leser, während für Bibliotheken und Informationseinrichtungen die kommunikative Funktion des Dokumentes im Vordergrund steht.⁶ Ebenso gilt für digitale Objekte, dass nicht primär die darin kodierten Informationen, sondern der semiotische Charakter der Daten von Interesse

⁴ Neben Buckland lassen sich zwei unabhängige Schulen feststellen, auf die hier nicht genauer eingegangen werden soll: zum einen ist dies die Kopenhagener Schule der ‘document theory’ (Hjørland 2007; Hjørland 2000; Lund 2009; Ørom 2007) und zum anderen die Arbeiten des französischen Kollektivs Roger T. Pédauque (2003; 2006; 2007; 2011).

⁵ Zitiert in Übersetzung nach Buckland (1997).

⁶ Zwar staubt ein Großteil von Dokumenten in Bibliotheken und Archiven im wahrsten Sinne des Wortes Jahrzehnte und länger vor sich hin, ohne dass eine Kommunikation stattfinden würde. Es ist jedoch auch nicht Aufgabe dieser Einrichtungen, ihre Bestände lediglich nach aktuellem Nutzwert inhaltlich auszuwerten, sondern die Möglichkeit einer wiederholten Kommunikation zwischen Autor und Leser über die Zeit hinweg aufrecht zu erhalten.

ist. So ist beispielsweise an einem Adressdatensatz weniger relevant, ob die Adresse korrekt angegeben ist, sondern dass es sich um einen Adressdatensatz handelt, der auf eine bestimmte Weise zu einem bestimmten Zweck zustande gekommen ist.

Fazit

Grundsätzlich lassen sich verschiedene Konzepte von Daten voneinander abgrenzen. Bei Diskursen und Aktivitäten um Daten sollte deshalb zunächst geklärt werden, welches Datenkonzept zugrunde liegt. Prinzipiell handelt es sich bei Daten um Unterschiede. Diese Unterschiede können im ersten Fall primär zwischen realen Vorgängen angenommen werden, so dass Daten als objektive Fakten auftreten. Im zweiten Fall spiegeln Daten vor allem Unterschiede zwischen verschiedenen Wahrnehmungen oder Beobachtungen wider, die auf mögliche oder tatsächliche Fakten verweisen. In der dritten und für die Bibliotheks- und Informationswissenschaft entscheidende Vorstellung von Daten spielen Fakten oder Inhalte nur eine sekundäre Rolle. Stattdessen sollten Daten vor Allem als Unterschiede zwischen Symbolen aufgefasst werden, die der Kommunikation dienen und je nach Nutzer unterschiedliche Informationen liefern können. Die kommunikativen Artefakte, als die Daten in diesem Fall auftreten, sind identisch mit digitalen Dokumenten.

Wie die passende Auseinandersetzung mit Daten bzw. digitalen Dokumenten genau auszusehen hat, steht noch nicht fest, da wir uns in einer Übergangsphase befinden (Dudek 2012). Der erste Schritt zu einem Verständnis von Dokumenten im rein Digitalen (Voß 2009) ist wahrscheinlich, sich mit den existierenden Formen, Funktionen und Praktiken von Daten vertraut zu machen. Für Bibliotheken und Archive bedeutet dies beispielsweise, Publikationen so zu nehmen, wie sie heute als Blogartikeln, Tweets, Slides, Videos u.v.a.m. auftreten — als Vorbild kann hier das Internet Archive dienen, das bereits seit 1996 digitale Publikationen verschiedener Art sammelt. Für die Auseinandersetzung mit Forschungsdaten kann die Fokussierung auf Daten bedeuten, dass es wichtiger ist, sich mit Daten- und Dateiformaten auszukennen und Forschungsdaten danach zu beurteilen, für was und vom wem sie in welchem Zusammenhang verwendet werden, statt welche Informationen in diesen digitalen Dokumenten “tatsächlich” enthalten sind.

Literatur

Ballsun-Stanton, Brian. 2010. Asking about Data: Experimental Philosophy of Information Technology. In: *5th International Conference on Computer Sciences and Convergence Information Technology*, 119–124.

---. 2012. Asking About Data: Exploring Different Realities of Data via the Social Data Flow Network Methodology. University of New South Wales.

Berners-Lee, Tim und Mark Fischetti. 1999. *Weaving the web: The original design and ultimate destiny of the world wide web by its inventor*. San Francisco: Harper.

Boell, Sebastian K. und Dubravka Cecez-Kecmanovic. 2011. Theorizing Information – From Signs to Sociomaterial Practices. In: *ACIS 2011 Proceedings*.

Briet, Suzanne. 1951. *Qu'est-ce que la documentation?*. Paris: Éditions documentaires, industrielles et techniques.

Buckland, Michael. 1997. What is a „document“?. *Journal of the American Society of Information Science (JASIST)* 48, Nr. 9 (September): 804–809. http://polaris.gseis.ucla.edu/gleazer/260_readings/Buckland.pdf.

---. 1998. What is a „digital document“?. *Document Numérique* 2, Nr. 2: 221–230.

Dudek, Sarah. 2012. Die Zukunft der Buchstaben in der alphanumerischen Gesellschaft. Text und Dokument unter digitalen Bedingungen. *Bibliothek Forschung und Praxis* 36, Nr. 2: 189–199. doi:10.1515/bfp-2012-0023, .

Floridi, Luciano. 2005. Is Information Meaningful Data?. *Philosophy and Phenomenological Research* 70, Nr. 2: 351–370. <http://philsci-archive.pitt.edu/archive/00002536/>.

---. 2009. Philosophical Conceptions of Information. In: *Formal Theories of Information*, 13–53. doi:10.1007/978-3-642-00659-3_2, .

---. 2010. *Information – a very short introduction*. Oxford University Press.

Gitelman, Lisa, Hrsg. 2013. *Raw data is an oxymoron*. MIT Press.

Gray, Robert L. 2003. Brief Historical Review of the Development of the Distinction Between Data and Information. In: *9th Americas Conference on Information Systems*, hg von. J. Ross und D. Galletta, 2843–2849. Tampa: Association for Information Systems.

Hjørland, Birger. 2000. Documents, memory institutions and information science. *Journal of Documentation* 56, Nr. 1 (Januar): 27–41.

---. 2007. Arguments for 'the bibliographical paradigm'. Some thoughts inspired by the new English edition of the UDC. *information research* 12, Nr. 4 (Oktober). <http://informationr.net/ir/12-4/colis/colis06.html>.

Jr., Earl H. McKinney und Charles J. Yoos II. 2010. Information About Information: A Taxonomy of Views. *MIS Quarterly* 34, Nr. 2: 329–344.

Kettinger, William J. und Yuan Li. 2010. The infological equation extended: towards conceptual clarity in the relationship between data, information and knowledge.. *European Journal of Information Systems* 19, Nr. 4: 409–421. doi:10.1057/ejis.2010.25, .

Lund, Niels Windfeld. 2009. Document theory. *ARIST* 43, Nr. 1: 1–55. doi:10.1002/aris.2009.1440430116, .

Naur, Peter. 1966. The science of datalogy. *Communications of the ACM* 9, Nr. 7 (Juli): 485.

Nelson, Ted. 2010. *Possiplex: Movies, Intellect, and Creative Control. My Computer Life and the Fight for Civilization*. 1 Aufl.. Mindful Press.

Otlet, Paul. 1934. *Traité de documentation. Le livre sur le livre. Théorie et pratique. IIB Publication*. Brussels: Editiones Mundaneum. <https://archive.ugent.be/handle/1854/5612>.

Pédauque, Roger T. 2003. Document: Form, Sign and Medium, As Reformulated for Electronic Documents. <http://archivesic.ccsd.cnrs.fr/docs/00/06/22/28/HTML/index.html>.

---. 2006. *Le document à la lumière du numérique*. Caen: C & F éditions.

---. 2007. *La redocumentarisation du monde*. Cepaduès.

---. 2011. Le web sous tensions. Februar. <https://espacestemps.co-ment.com/text/vsMAqHUTfli/view/>.

Rayward, W. Boyd. 1994. Visions of Xanadu: Paul Otlet (1868-1944) and Hypertext. *Journal of the American Society for Information Science* 45, Nr. 4 (Mai): 235–250. http://people.lis.uiuc.edu/~wrayward/Visions_of_Xanadu_JASIS.pdf.

Shannon, Claude Elwood. 1948. A mathematical theory of communication. *Bell Systems Technical Journal* 27: 379–423.

Voß, Jakob. 2009. Zur Neubestimmung des Dokumentenbegriffs im rein Digitalen. *Libreas - Library Ideas* 15 (Oktober). <http://libreas.eu/ausgabe15/texte/002.htm>.

---. 2011. Revealing digital documents. Concealed structures in data. 1105.5832.

---. 2013. Describing data patterns. A general deconstruction of metadata standards. Berlin: Humboldt University.

Walther, Marcel, Marleen Burger, Maxi Kindling, Lisa Liebenau, Claudia Lienhard, Svantje Lilienthal, Paulina Plewka, u. a. 2013. *Forschungsdatenmanagement an Hochschulen – Internationaler Überblick und Aspekte eines Konzepts für die Humboldt-Universität zu Berlin*.

Ørom, Anders. 2007. The concept of information versus the concept of document. In: *Document (re)turn. Contributions from a research field in transition*, hg von. Roswitha Skare, Niels Windfeld Lund, und Andreas Vårheim, 53–72. Frankfurt: Peter Lang.