
Themenheft Nr. 37: Medienpädagogik als Schlüsseldisziplin in einer mediatisierten Welt. Perspektiven aus Theorie, Empirie und Praxis
Herausgegeben von Henrike Friedrichs-Liesenkötter, Lara Gerhardts,
Anna-Maria Kamin und Sonja Kröger

«Big Data» aus der Perspektive von Informatischer Bildung und Medienpädagogik Bildung in einer datifizierten Gesellschaft

Johannes Magenheim

Zusammenfassung

Mit der Verbreitung und scheinbaren «Allgegenwart» von IuK-Technologien in vielen gesellschaftlichen Bereichen geht das Erzeugen von grossen Datenmengen einher. Diese «Big Data» sind vielfach unstrukturiert, auf viele Quellen verteilt, teilweise unvollständig und zumeist nicht miteinander verknüpft. Sie können von verschiedenen Interessensgruppen für je spezifische Zwecke ausgewertet werden. Die damit verbundenen gesellschaftlichen Auswirkungen zeigen sich u.a. auf sozialer, ökonomischer, politischer, ethischer und kultureller Ebene. Da Digitale Medien und Informatiksysteme bei der Erzeugung, Verarbeitung und Nutzung von «Big Data» eine zentrale Rolle spielen und sie, z.B. in Form von «Social Media», einen hohen Bezug zur Alltagswelt von Jugendlichen haben, kommt ihnen auch in der Medienpädagogik und der informatischen Bildung eine erhebliche Bedeutung zu. Der Beitrag setzt sich mit gesellschaftlichen Auswirkungen von «Big Data» auseinander und versucht aufzuzeigen, welche Aufgaben daraus für die Medienpädagogik und die informatische Bildung erwachsen. Es werden gemeinsame, sich ergänzende und spezifische Aufgabenfelder für beide Bildungsbereiche aufgezeigt und die Notwendigkeit einer interdisziplinären multiperspektivischen Sicht auf «Big Data» im Unterricht begründet. Für einen derartigen Unterricht

stehen eine Reihe von geeigneten Softwaretools zur Verfügung, die im Informatikunterricht aber auch in anderen Fächern oder in fachübergreifenden Projekten genutzt werden können.

«Big Data» from the perspective of computer science education and media pedagogy

Abstract

The ubiquitous presence of information and communication technologies in many areas of society is accompanied by the generation of large amounts of data. This «Big Data» is often unstructured, distributed over many sources, partly incomplete and mostly not linked with each other. Different interest groups can evaluate them for specific purposes. The associated social impact can be seen on a social, economic, political, ethical, and cultural level. Digital media and Informatic Systems play an essential role in the generation, processing, and use of «Big Data». For example, «Social Media» have high relevance to the everyday life of young people. Therefore, «Big Data» also pose a significant educational challenge for Media Pedagogy and Computer Science Education. The article deals with the social effects of «Big Data» and tries to reveal which tasks arise for media pedagogy and computer science education. Joint, complementing, and specific tasks for both educational fields are pointed out, and the necessity of an interdisciplinary multi-perspective view of «Big Data» in education is justified. Several suitable software tools are available for such learning and teaching, which can be used in computer science lessons, in other subjects or interdisciplinary projects.

Einleitung

In einer durch zunehmende Digitalisierung geprägten Gesellschaft entstehen, bedingt durch vielfältige Kommunikations-, Interaktions- und Produktionsprozesse unterschiedlichster Art, grosse, verteilte, meist voneinander unabhängige Datensammlungen. Im Gegensatz zu strukturierten Datenbeständen, wie wir sie mit den Datenbanken aus der traditionellen

Informatik kennen, sind diese Daten völlig unterschiedlich oder gar nicht strukturiert und zum Teil auch unvollständig. Diese Datenbestände werden oft mit dem Oberbegriff ‹Big Data› (BD) bezeichnet. ‹Big Data› und die zu ihrer Generierung und weiteren Verarbeitung verwendeten Algorithmen besitzen einen Doppelcharakter. Einerseits bilden sie Vorgänge des realen Lebens in digitaler Form ab und stellen so eine Projektion von Realität in die virtuelle Welt dar. Andererseits können sie auf vielfältige Weise mittels geeigneter Informatiksysteme von Nutzergruppen mit spezifischen Nutzungsinteressen auch dazu verwendet werden, um Prozesse des realen Lebens zu beeinflussen und zu steuern. Am Beispiel der Autofahrt einer vierköpfigen Familie in den Urlaub (in naher Zukunft) soll dies anhand eines kleinen Ausschnitts von sozialer und technischer Realität verdeutlicht werden:

Die Frau sitzt beobachtend am Steuer, das Auto wird weitgehend autonom durch den Verkehr gesteuert und tauscht sich mit den anderen Fahrzeugen über Fahrgeschwindigkeit und Abstand aus. Ein Verkehrsleitsystem beeinflusst die Geschwindigkeit der Fahrzeugkonvois, um Staus zu minimieren. Die Fahrtroute wird in einem grossen Display neben dem Lenkrad angezeigt. Je nachdem, in welchem Modus sich dieses Informationssystem befindet, können verschiedene Audio-Medien wiedergegeben, Informationen über aktuelle Fahrdaten angezeigt und Fahrparameter verändert werden. Bei Bedarf liest das Info-System eingehende E-Mails und Instant-Messages vor oder stellt mittels Sprachbedienung Telefonverbindungen zu den im System gespeicherten Kontakten der Fahrerin her. Die Frau hat den Fahrstil Economy am Fahrzeug eingestellt, da dieser Voraussetzung für einen günstigen KfZ-Versicherungstarif ist. Ausserdem wird mittels Gesichtserkennung für die Versicherung festgestellt, ob sie selbst am Steuer sitzt. Die Fahrdaten des Fahrzeugs werden an Versicherung, Autohersteller und eine Mautagentur übermittelt, um die CO₂-bezogene Umweltsteuer für die Fahrt zu berechnen. Für das Aufladen des Hybridfahrzeugs mit Strom werden in Abhängigkeit von den verfügbaren Ladestationen an der Strecke und dem Speicherbedarf des Fahrzeugs sowie der Lastverteilung im nationalen Stromnetz die Pausen- und Standzeiten des Fahrzeugs auf der Fahrt in den Urlaub ermittelt. Im Auto ist ein WLAN-Hotspot aktiviert. Der Mann liest auf dem Tablet die elektronische Ausgabe einer

Zeitung, tätigt eine Überweisung und schaut sich den Wetterbericht für den Urlaubsort an. Dabei bemerkt er, dass ihm am Rande der Webseite verschiedene Werbeangebote für Sportschuhe gemacht werden. Nach solchen hatte er in den vergangenen Tagen im Internet immer wieder gesucht. Die beiden Kinder nutzen ihre Tablets mit Bluetooth-Verbindung zu ihren Headsets. Sie versenden und empfangen Posts und von Ihnen modifizierte Fotos in verschiedene Social Media, hören von einer Plattform von dieser entsprechend ihrem Geschmack empfohlene Musik und schauen sich die Top-gerankten Videoclips an.

Zur Pause, während des Halts an der Ladestation, betätigen sich die Familienmitglieder sportlich. Die Fitnessstracker an den Armen der Familienmitglieder speichern Puls und Laufstrecke, die per Algorithmus zu einem täglichen Fitnessreport zusammengefasst und der an die Krankenkasse der Familie übermittelt wird. Dies ist die Bedingung, um den günstigen Krankenversicherungstarif zu behalten.

Zugegeben, eine (hoffentlich) noch nicht ganz reale Alltagssituation, obwohl die genannten technischen Features allesamt bereits grundsätzlich verfügbar sind und viele schon jetzt zum medialen Alltag gehören. Das Beispiel zeigt, wie stark die Digitalisierung und ‹Big Data› nicht nur technische Abläufe betreffen, sondern wie stark die Quantifizierung des Sozialen (Mau 2018) allgegenwärtig ist und die Steuerung von Sozialverhalten ermöglicht. Die bei der Erfassung und Auswertung von grossen Datensammlungen angewandten, technischen, informatischen und mathematischen Methoden werden vornehmlich der Wissenschaftsdisziplin ‹Data Science› (DS) zugeordnet.

Dieses Beispiel verdeutlicht zugleich auch wesentliche Charakteristika von ‹Big Data›:¹

- Der Zugriff auf und die Analyse von ‹Big Data› stellen aus informatischer Perspektive hohe Anforderungen an Auswertungsmethodik und informationstechnische Infrastruktur.
- Einige Anwendungen können dabei auf gespeicherte Daten im Batchbetrieb zugreifen und sind nicht zeitkritisch, während andere eine kontinuierliche Echtzeitverarbeitung der Daten mit zum Teil hoher Geschwindigkeit erfordern.

1 (Alharthi, Krotov, und Bowman 2017; Yaoxue Zhang u.a. 2017; Chen, Mao, und Liu 2014; De Mauro, Greco, und Grimaldi 2015).

- Aus der in den Daten enthaltenen, z.T. versteckten Information kann Wissen erzeugt werden, das für Individuen oder Organisationen von hohem Nutzen und Wert sein kann. Es geht bei der Analyse der Daten darum, ihre Bedeutung für einen spezifischen Anwendungskontext zu verstehen. Dies kann aufgrund daraus abgeleiteter Handlungsstrategien zu ökonomischem Gewinn, aber auch zu politischen oder sozialen Vorteilen für diejenigen führen, die Zugang zu diesem Wissen erlangen.
- Da die Daten aus unterschiedlichen Quellen stammen, müssen sie in geeigneter Weise zusammengeführt und ausgewertet sowie die Ergebnisse im Hinblick auf ihren Wahrheitsgehalt und die adäquate Abbildung der Realität überprüft werden. Nur auf diese Weise kann eine neue Qualität der aggregierten Daten und damit ein qualitativ höherer Nutzen für die Anwender erzielt werden.
- In vielen Fällen werden persönliche Daten der mit den Informatiksystemen interagierenden Personen verdeckt oder offen erfasst, sodass hohe technische Standards bei der Datensicherheit und rechtliche Normen beim Datenschutz einzuhalten sind.

Erfassung und Auswertung von BIG Data und deren gesellschaftliche Auswirkungen

Das oben geschilderte Beispiel zeigt, dass die Nutzung digitaler Medien, je nach Betrachtungsperspektive auf den sozialen Kontext in der sie stattfindet, von sehr unterschiedlicher Komplexität sein kann. Das reicht vom Umgang einer Person mit einem Tablet oder einem Display im Fahrzeug bis hin zu den komplexen Informatiksystemen, die hinter den auf den Geräten angebotenen Diensten von Apps oder der Fahrzeugsteuerung per 5G-Technik stecken.

Um ein BD-Anwendungsszenario zu analysieren und seine gesellschaftlichen Auswirkungen zu beurteilen, sollten verschiedene Dimensionen betrachtet und miteinander kombiniert werden. Auf diese Weise wird ein multiperspektivischer Blick auf einen ausgewählten BD-Gegenstandsbereich eröffnet.

Die Analyse des Einsatzes und der sozialen Auswirkungen von ⟨Big Data⟩ kann strukturiert werden nach:

– *Komplexität des sozialen Kontextes*

Je nach Komplexität des sozialen Kontextes des Anwendungsszenarios ergeben sich Auswirkungen des Einsatzes von DS-Methoden auf individueller Ebene, auf der Ebene von Gruppen, einer Organisation / eines Unternehmens, auf gesamtgesellschaftlicher oder sogar auf internationaler oder globaler Ebene. Diese Auswirkungen sind oft nur im interdependenten Zusammenwirken der verschiedenen Ebenen und zueinander komplementär zu verstehen.

– *den gesellschaftlichen Anwendungsbereichen*

Es gibt eine Vielzahl von gesellschaftlichen Anwendungsbereichen, in denen Methoden und Konzepte von BD zum Einsatz kommen. Als wichtige Bereiche können angesehen werden:

- *Ökonomie*: z.B. Prognose von wirtschaftlichen Entwicklungen;
- *Energie*: z.B. Balance von Erzeugung und Verbrauch im Netz;
- *Umwelt/Ökologie*: z.B. Wettervorhersage, Klimaprognosen;
- *Katastrophen-Management*: Datenanalyse zur Risikominderung;
- *Gesundheitswesen*: Diagnose und Ursachenforschung von Krankheiten;
- *Politik*: z.B. Datenanalyse für politisch-strategische Entscheidungen;
- *E-Government*: z.B. Automated Decision Making mit ADM-Algorithmen;
- *Bildungswesen*: z.B. Learning Analytics;
- *Verkehr*: z.B. autonomes Fahren;
- *Smart Home, Smart Cities*: z.B. Steuerung von Service-Angeboten, Verkehr;
- *Kultur*: z.B. Produktion künstlicher von realen nicht zu unterscheidenden Filmen;
- *Social Media*: z.B. Steuerung von Konsumentenverhalten, politische Meinungsbildung.

– *nach der Art der sozialen Auswirkungen*

- *Ökonomisch*: Arbeitsplätze, Entscheidung über wirtschaftliche Strategien.
- *Politisch*: Politischer Missbrauch (Überwachung) versus verbesserte Grundlagen für Entscheidungsfindung, politische Partizipation.

- *Sozial und ethisch*: Partizipation, Gleichheit, Vertrauen, Diskriminierung (Exklusion, Inklusion).
- *Rechtlich*: Datenschutz, Urheberrecht, Vertrauen und Zugangsrechte, Datensicherheit.

Die auf breiter Basis erhobenen persönlichen Daten können zu deren Missbrauch und zur Verletzung von Persönlichkeitsrechten führen. Dies gilt auch für den Umgang mit Urheberrechten. Hier ist eine Balance zwischen Freiheit der Meinungsäußerung im Netz, Beachtung von Persönlichkeitsrechten und Verhinderung politischer Zensur im Internet zu wahren.

Ethische, soziale und rechtliche Aspekte wirft auch der Einsatz von ADM-Algorithmen auf.

Es stellen sich Fragen nach ihrer informatischen Realisierung (Funktion der Algorithmen und deren Semantik und Syntaktik, Qualität der herangezogenen Datenquellen, Art des Trainings von verwendeten neuronalen Netzen etc.) und nach der Transparenz und logischen Nachvollziehbarkeit der dadurch induzierten Entscheidungen. Oftmals müssen die in Bruchteilen von Sekunden, ggf. ohne die Möglichkeit menschlicher Intervention, getroffen werden, wie z.B. bei Kollisionsvermeidungsentscheidungen beim autonomen Fahren. Aber auch, wenn es sich nicht um ein Echtzeitproblem handelt, etwa bei der Gewährung eines Kredits oder der Frage nach Inklusion oder Exklusion von Personengruppen bei einer Versicherung, spielen ethische, soziale, rechtliche und ökonomische Aspekte eine wichtige Rolle (Bigdatasoc 2019; Cuquet u.a. 2017; Lepri u.a. 2017; Mantelero 2018; Mittelstadt u.a. 2016; EU Expertengruppe für künstliche Intelligenz 2019; Kolmar und Binswanger 2019; Heckmann 2019).

Als ein weiteres Beispiel für die wechselseitige Abhängigkeit von sozialer Komplexität, gesellschaftlichem Anwendungsbereich sowie Art der sozialen Folgewirkung sei hier das ‹Micro-Targeting› erwähnt. Durch die systematische Analyse von Daten aus sozialen Netzwerken, sowie zum Wählerverhalten und der Meinungsforschung mittels DS-Methoden werden zielgruppenbezogene Werbemaßnahmen zwecks politischer Meinungsbeeinflussung ermöglicht. Dies kann z.B. durch Posts in Social Media geschehen, die nur für eine bestimmte Zielgruppe sichtbar sind und

deren Meinung in der ‹Filter Bubble, Echo Chamber› noch verstärkt. Auf diese Weise können auf nationaler Ebene Wählerverhalten beeinflusst und Wahlen entschieden werden. Dies hat z.B. zu einer Milliardenstrafe für Facebook geführt. Facebook-Daten waren von der Firma Cambridge Analytics zu gezielten Micro-Targeting Aktionen im amerikanischen Wahlkampf verwendet worden (dpa 2019).

Insgesamt kann man aus den genannten Beispielen die Schlussfolgerung ziehen, dass es Aufgabe der Medienpädagogik und der informatischen Bildung ist, Individuen zu befähigen, die ihnen offerierten medialen Angebote auf ihren Wahrheitsgehalt und Realitätsbezug hin kritisch zu überprüfen. Ausserdem sollten Personen zu komplexem, vernetztem Denken befähigt werden, um sich als mündige Bürger und Bürgerinnen einen multi-perspektivischen Blick auf digitale Medien, ‹Big Data› und deren gesellschaftliche Folgewirkungen zu erschliessen.

Multiple Sichten auf Big Data und Digitale Medien in der Bildung

Die technologische und gesellschaftliche Komplexität des Themas ‹Big Data› zeigt, dass es im Sinne von Allgemeinbildung als epochaltypisches Schlüsselproblem (Klafki 1993) ein wichtiger Unterrichtsgegenstand sein kann. Die Komplexität und die Diversität der gesellschaftlichen Erscheinungsformen dieses Gegenstandsbereichs stellt die schulische Bildung, insbesondere die Medienbildung und die informatische Bildung, vor grosse Herausforderungen.

Eine Studie zur Bildung für und über ‹Big Data› (Gapski, Tekster, und Elias 2018) plädiert für einen multi-perspektivischen, interdisziplinären Blick auf das Thema ‹Big Data›. Um dies zu ermöglichen, sollten politische Rahmendiskurse (wirtschafts- und arbeitsmarktpolitisch, bildungspolitisch, zivilgesellschaftlich-digitalpolitischer) geführt werden. Der für ‹Big Data› wichtige Bildungsdiskurs wird hinsichtlich qualifikatorischer, informatischer, medienpädagogischer, medien-kulturell-ästhetischer, politischer, datenschutzorientierter, verbraucherpolitischer und ethischer Aspekte als notwendig erachtet.

Im Hinblick auf die Medienpädagogik und die informatische Bildung zeigt das Eingangsbeispiel, wie unscharf der Begriff ‹Digitale Medien› geworden ist und wie stark er, zumindest aus der Perspektive der Informatik, mit dem Begriff des digitalen Artefakts bzw. des sozio-technischen Informatiksystems verknüpft ist (Schulte u.a. 2017). Der Zugriff auf eine App mittels Tablets hat vordergründig zunächst einmal mit Mensch-Maschine-Interaktion und der sachgemässen Bedienung des digitalen Medienobjekts zu tun. Gleichzeitig werden jedoch im Hintergrund, für die Nutzenden nicht sichtbar, auf informationstechnischer Ebene eine Reihe von Interaktionen zwischen diesem digitalen Artefakt und komplexeren Informatiksystemen in Gang gesetzt. Das können funktionsbezogene Daten der App sein, die zur weiteren Verarbeitung der Nutzereingaben auf einen entfernten Server gesendet und nach Verarbeitung von dort sekundenschnell wieder zur App zurück übertragen werden. Das können aber auch private nutzerbezogene Daten sein, die von einer App, verdeckt und teilweise illegal, von anderen auf dem mobilen Endgerät installierten Apps abgeschöpft und übermittelt werden können. Hinzu kommt, dass viele Apps, wie z.B. die zum Streamen von Audio- und Videoangeboten, eine Social-Media-Komponente besitzen, die es den Nutzenden ermöglicht, ihre aktuellen Aktivitäten auf einer App per Social Media mit anderen zu teilen. Daher kann der Umgang mit digitalen Medien und die durch sie erzeugten Datenmengen, ergänzend zu den oben beschriebenen Analyseebenen in Bezug auf die sozialen Auswirkungen von ‹Big Data› um weitere ergänzt werden. Dies gilt insbesondere auch für den Bereich der Nutzung von Digitalen Medien im Bereich der Bildung:

- Art und Intention der Nutzung des digitalen Medienobjekts (Medium zum Lernen und/oder digitales Tool)
- Mensch-Maschine Interaktion mit dem digitalen Medienobjekt
- Induzierte technische Kommunikation zwischen digitalen Objekten und Informatiksystemen
- Induzierte soziale Kommunikation in assoziierten sozialen Netzwerken
- Generierte Datenströme und Datensammlungen

Diese informatisch geprägte Sicht auf die Nutzung digitaler Medien erfolgt in Anlehnung an eine systemtheoretische Auffassung von Technik

(Ropohl 2009), die ein sozio-technisches Informatiksystem als die Einheit von Software, Hardware und assoziiertem sozialen Handlungssystem von Personen begreift, die mit dem technischen Teil des Systems und miteinander in einem entsprechenden Anwendungskontext des Systems interagieren (Magenheim 2008). Aus informatischer Perspektive gilt es dabei zwischen den nach aussen sichtbaren Funktionen des Informatiksystems (z.B. grafische Benutzungsoberfläche) und seiner damit zusammenhängenden zunächst nicht sichtbaren inneren Struktur (z.B. Algorithmen) zu unterscheiden. Die Offenlegung (Dekonstruktion) dieser vielfältigen Beziehungen, die das soziale System der interagierenden Nutzenden einschliesst, und Informatiksysteme so als hybride Systeme begreift (Budde und Schulte 2018), ist eine wichtige Aufgabe der informatischen Bildung auch bei der Analyse digitaler Medien.

«Big Data» können bei der Nutzung von Informatiksystemen in vielfältiger Weise entstehen (siehe Abschnitt 2). Insbesondere bei der Nutzung von Social Media entstehen auf diese Weise Artefakt-Aktor-Netzwerke (AAN), die die von den Akteuren mittels digitaler Medien erzeugten digitalen Objekte sowohl auf technischer als auch auf sozialer Ebene miteinander verbinden (Riss u.a. 2011).

Für einen multiperspektivischen Blick auf digitale Medien, der insbesondere auch die informatischen Grundlagen der digitalen Medien berücksichtigt, wird auch seitens Vertretern und Vertreterinnen der Medienpädagogik geworben. Zorn (2011, 180) plädiert dafür, Digitale Medien «weniger aus einer medientheoretischen Perspektive, sondern stärker auch aus einer softwaretheoretischen Perspektive» zu betrachten und zu verstehen. Herzig (2016, 74) stellt fest, dass sich für einen interdisziplinären Bereich der Medienbildung «hinreichend «Spuren» in der Medienpädagogik und in der Informatik – ergänzend auch in der Medienwissenschaft – finden» lassen. Die Zusammenhänge zeigen sich auf der technischen Ebene der Maschine (Bearbeitung von sinnentleerten Formalismen), auf der symbolischen Ebene (Software, Dualität von Semiotik und Technik, Modellierung) und auf der kognitiven Ebene der involvierten Personen (Denk- und Handlungsstrukturen, Interpretation von Daten, Re-kontextualisierung) sowie in den Wechselwirkungen zwischen diesen (vgl. auch Herzig 2016, 69f.; Eder, Mikat, und Tillmann 2017; Tulodziecki 2016).

Der Diskurs zwischen Medienpädagogik und Informatischer Bildung hat, nicht zuletzt in Erkenntnis dieser Zusammenhänge, in jüngster Zeit zu gemeinsamen publizierten Dokumenten geführt (Dagstuhl-Erklärung: Brinda u.a. 2016; Frankfurt-Dreieck: Brinda u.a. 2019), die die Grundlage für einen interdisziplinären Ansatz von Medienbildung und informatischer Bildung schaffen könnten.

Jedem Betrachtungsgegenstand dieses interdisziplinären Ansatzes sollen Analyse, Reflexion und Gestaltung zuteil werden. Als Betrachtungsperspektiven sind vorgesehen:

- *Technologisch-mediale Perspektive*: «Hinterfragen und Reflektieren der den Phänomenen und Artefakten der durch Digitalisierung geprägten Welt zugrundeliegenden Strukturen und deren Funktionsweisen sowie eine Befähigung zur (Mit-) Gestaltung solcher Artefakte und Phänomene».
- *Gesellschaftlich-kulturelle Perspektive*: «Wechselwirkungen zwischen Individuen, Gesellschaft und digitalen Systemen vor dem Hintergrund der Mediatisierung und des digitalen Wandels» analysieren und reflektieren.
- *Interaktionsperspektive*: Im Mittelpunkt der Betrachtung stehen Menschen und «die Aspekte Nutzung, Handlung und Subjektivierung». Wichtige zu beantwortende Fragen sind, wie Menschen «welche digitalen Medien und Systeme warum und wozu nutzen, inwiefern sie am digitalen Wandel teilhaben und ihn mitgestalten (können) sowie wie sie sich als handlungsfähige Subjekte konstituieren» (vgl. Brinda u.a. 2019).

Das Thema «Big Data» wäre mit seinen vielfältigen interdisziplinären Bezügen im Hinblick auf gesellschaftliche Anwendungsfelder, die daraus resultierenden sozialen Folgewirkungen sowie den informatischen und medialen Aspekten als Betrachtungsgegenstand im Sinne des mit dem Frankfurter Dreieck beschriebenen Ansatzes hervorragend geeignet.

«Big Data» als Herausforderung für die Medienpädagogik

Die Vorgaben bezüglich des Umfangs dieses Artikels lassen es nicht zu, die aktuelle medienpädagogische Diskussion zum Thema «Big Data Analytics» umfassender darzustellen. Zusammenfassend ist zu konstatieren, dass digitale Medien und auch das Thema «Big Data» aus der Sicht der Medienpädagogik «eine wichtige Rolle als Lerngegenstand und Lernmittel in Bildungsprozessen einnehmen» (Albrecht und Revermann 2016, 50). Viele Autoren und Autorinnen sind zudem der Auffassung, dass sich die Aufgabe der Medienpädagogik nicht nur an der Vermittlung von Kompetenzen orientieren darf, sondern, dass sich bei den Lernenden auch eine Handlungspraxis herausbilden sollte, die auf der «Veränderung von Welt- und Selbstbezügen» beruht (ebd.). Dieser auf praktische Handlungsfähigkeit im Umgang mit «Big Data» abzielende Aspekt von Medienkompetenz wird z.B. auch von Brüggem (2016, 20) betont. Es gehe um «Wissen und Bewerten als Grundlage für Orientierung und gesellschaftliche Handlungsfähigkeit in einer mediatisierten Gesellschaft».

Im Zusammenhang mit «Big Data Analytics» werden als Aufgaben der Medienpädagogik angesehen «[...] Ohnmachtserfahrung als Ausgangspunkt ernst zu nehmen und in der Arbeit produktiv zu wenden [...] Hintergründe zu verstehen und eigene Vorstellungen zu entwickeln [...]. bei der Suche nach Lösungen den Fokus über die individuelle Selbstverantwortung auszuweiten» (ebd., 20f.).

Beim Thema «Big Data» als Lerngegenstand und Lernmittel in der Medienbildung werden einerseits von verschiedenen Publizierenden die interdisziplinären Bezüge betont, andererseits darauf verwiesen, dass schon bestehende medienpädagogische Konzepte lediglich im Hinblick auf den «Betrachtungsgegenstand» «Big Data» hin spezifiziert werden müssen.

Dementsprechend verweist Tulodziecki (2016b) darauf,

«dass die strukturelle Frage der Steuerung, Manipulation und Kontrolle individuellen Verhaltens und gesellschaftlicher Entwicklungen durch Medien schon einen wichtigen Stellenwert in der Medienpädagogik hatte, bevor das Thema Big Data relevant wurde. So spielt z.B. das Durchschauen und die Kritik von technologischen, rechtlichen, ökonomischen, institutionellen und weiteren gesellschaftlichen

Bedingungen der Medienproduktion und Medienverbreitung in verschiedenen Ansätzen zur Medienkompetenz eine wichtige Rolle » (Tulodziecki 2016a, 87).

In einem der Kompetenzmodelle wird sogar explizit auf die Bedeutung von sozio-technischen Systemen als Komponente von Medienkompetenz hingewiesen (Schiersmann, Busse, und Krause 2002, 40): «Nutzung und Gestaltung technischer Systeme, Gestaltung sozio-technischer Systeme, Kundige Kritik von Medien und IuK-Technologien». Aßmann u. a. (2017) beschreiben mit etwas anderen, aber inhaltlich ähnlichen Kategorien, wie in diesem Beitrag in Abschnitt 2 geschehen, wichtige Problemfelder von «Big Data», mit denen sich die Medienpädagogik auseinandersetzen sollte: Privatheit und Öffentlichkeit, Informationelle Selbstbestimmung, Datenerhebungs- und Datenauswertungspraktiken, Datenkapitalismus, Menschliche Berechenbarkeit und Mensch-Maschine Verhältnis. Als «Schlüsselfunktion und Aufgabenfelder der Medienpädagogik» werden dabei beschrieben:

1. Abstraktheit, Unsichtbarkeit und Komplexität in Anschaulichkeit übersetzen
2. Digitale Infrastrukturen durchschauen und demokratisch mitgestalten
3. Meinungsvielfalt einfordern und Diskriminierungen anzeigen
4. Die digitale Selbstbestimmung fördern
5. Produktive und gesellschaftlich wünschenswerte Nutzungsformen ermöglichen.

Mit diesem Ansatz zum Umgang mit «Big Data» in der Medienpädagogik werden die politisch-sozialen, ethischen und ökonomischen Aspekte, die Medienpraxis der handelnden Subjekte sowie auch die informationstechnischen Aspekte von «Big Data» adressiert. Die Autoren und Autorinnen fordern, dass die Medienpädagogik, «das sich verändernde Mensch-Maschine-Verhältnis reflektieren» müsse. Es gehe «sowohl darum, das Zusammenwirken analog-digitaler Schnittstellen zu verstehen, als auch die Charakteristika und Funktionen, die sich aus Vernetzung und Analyse ergeben. Dazu müssen sowohl technische, informatikbezogene praktische Zugänge geschaffen werden als auch Zugänge, die die ethischen Konsequenzen in den Blick nehmen» (Aßmann u.a. 2017, 190f.).

«Big Data» aus der Perspektive informatischer Bildung

Bezüge zur Fachwissenschaft Informatik und zu Data Science

In neueren Diskussionen und Curricula der Informatikdidaktik auf nationaler und internationaler Ebene spielt das Thema «Big Data» und «Data Science» sowohl hinsichtlich seiner informatischen, mathematischen und informationstechnischen Grundlagen als auch bezüglich seiner gesellschaftlichen Auswirkungen eine wichtige Rolle (siehe z.B. Grillenberger und Romeike 2018; Ridsdale u.a. 2015; Magenheim und Schulte 2020; Biehler und Schulte 2018; Royal Society 2017; Computer Science Teachers Association 2017).

Die Bedeutung des Themenbereichs für die informatische Bildung wird dabei einerseits mit der zunehmenden Relevanz von Data Science Methoden in der Bezugswissenschaft Informatik begründet. Andererseits werden die sozialen, ökonomischen, politischen und kulturellen Auswirkungen als wichtige Argumente angesehen, um sich mit dem Thema «Big Data» im Informatikunterricht zu beschäftigen.

Die Herausforderungen, die mit der Analyse und Verarbeitung von «Big Data» verbunden sind, kann man zusammenfassend mit den «6V» charakterisieren: *Volume* (Umfang der Datensammlung), *Velocity* (Geschwindigkeit des Datenaufkommens), *Variety* (Vielfalt der Datenstrukturen und -quellen, wie Text, Symbole, Audio, Bilder, Videos, sequentielle Datenströme von Sensoren), *Variability* (Veränderung der Datenlage und der Datenmodelle), *Veracity* (Authenzität, Vertrauenswürdigkeit), *Visualization* (Problemangemessene Darstellung der entdeckten Zusammenhänge), *Value* (Wert der gewonnenen Einsichten (statistisch überprüfte Hypothesen), prädiktive Aussagekraft) (vgl. z.B. Sivarajah u.a. 2017, 265).

Die zu diesen Herausforderungen adäquaten informatischen Methoden wie «Machine Learning» oder «Deep Learning» arbeiten mit Algorithmen, die Daten z.B. hierarchisch in Entscheidungsbäumen strukturieren oder Datencluster im Hinblick auf Ähnlichkeiten ggf. unter Berücksichtigung von Wahrscheinlichkeitswerten berechnen.

Mit Hilfe von neuronalen Netzen können bereits bekannte Muster in Datenmengen identifiziert (Gesichtserkennung, Diagnose von

Krankheiten) oder neue Muster entdeckt sowie weitere Entwicklungen aus bisherigen Verläufen der Datengenerierung prognostiziert werden.

Auf schulischer Ebene sind derartige professionelle Verfahren nur exemplarisch in vereinfachter und grundlegender Form behandelbar. So ist es an einfachen Beispielen möglich, Grundkonzepte des maschinellen Lernens und der Arbeitsweise von neuronalen Netzen zu thematisieren. Zudem müssen zum Verständnis dieser DS-Methoden auch Grundlagen der Statistik und der Wahrscheinlichkeitsrechnung zusätzlich zur Algorithmik in den Informatikunterricht mit einbezogen werden (Biehler und Schulte 2018; Magenheim 2018).

Bei der Suche nach in diesem Sinne zu erweiternden Kompetenzmodellen und Curricula, die das Thema BD in der schulischen informatischen Bildung berücksichtigen, orientieren sich viele Autoren am Data Life Cycle (DLC) (vgl. Biehler und Schulte 2018; Faundeen 2012; Grillenberger und Romeike 2018; Magenheim 2018) und betrachten somit ‹Big Data› primär aus der Perspektive der Daten und den Phasen ihrer Bearbeitung. Eine andere Perspektive, die sich mit der ersten z.T. überschneidet, ist die Anlehnung an professionelle Prozessmodelle des ‹Data Mining› und des ‹Data Managements›, z.B. dem CRISP-DM Modell (De Veaux u.a. 2017). Hier wird die Bearbeitung und Analyse von ‹Big Data› aus der Perspektive der Datennutzung in Organisationen und Unternehmen betrachtet. Dem *Data Life Cycle* werden die folgenden Phasen zugeordnet.

- *Planen*: Beschreibung der Daten, die gesammelt werden sollen, und wie die Daten während ihrer gesamten Lebensdauer verwaltet und verfügbar gemacht werden
- *Sammeln*: Daten erfassen
- *Sicherstellen*: Qualität der Daten prüfen, ggf. Daten bereinigen
- *Beschreiben*: Daten werden mit den entsprechenden Metadatenstandards genau und gründlich beschrieben
- *Bewahren*: Daten in einem geeigneten Langzeitarchiv speichern
- *Entdecken*: Potenziell nützliche Daten werden lokalisiert und erhalten
- *Integrieren*: Daten aus unterschiedlichen Quellen zu einem homogenen analysierbaren Datensatz zusammenfassen
- *Analysieren*: Daten werden analysiert, um relevante Informationen zu gewinnen

Andere Konzepte haben zusätzlich weitere Phasen ergänzt, darunter einige technische Aspekte des Datenhandlings (siehe z.B. Grillenberger und Romeike 2018).

- *Modellieren*: Strukturieren der Daten, sodass sie gespeichert und effizient abgerufen werden können.
- *Implementieren*: Implementieren des Datenmodells und Speichern der digitalen Daten.
- *Visualisieren*: visuelle Darstellung und Bearbeitung von gewonnener Information, um sie besser verständlich zu machen.
- *Bewerten*: z.B. die Analyseergebnisse und die Qualität der Daten, auf denen sie basieren.
- *Teilen*: Daten (ursprüngliche oder aggregierte) und/oder die Ergebnisse der Analyse mit anderen teilen.

Diesen Phasen des *Data Life Cycles* können im Informatikunterricht methodische Phasen eines Data Science Unterrichtsprojekts zugeordnet werden, in dem in einem für Lernende angemessenen Anspruchsniveau Probleme von *Big Data* in einem spezifischen gesellschaftlichen Anwendungsbereich thematisiert werden. Zugleich können die mit den Projektphasen verbundenen Tätigkeiten auch als Kompetenzanforderungen bzw. Aspekte eines Kompetenzmodells zum Umgang mit *Big Data* verstanden werden. Diese aus dem DLC abgeleiteten Kompetenzanforderungen beinhalten Komponenten, die besonders bei der Datenmodellierung, der Datenintegration und der Datenanalyse über bisherige z.B. in den Bildungsstandards (GI – Gesellschaft für Informatik e.V. 2008; 2016) formulierten Datenbezogenen Kompetenzanforderungen hinausgehen. Die Prozessperspektive des CRISP-DM Modells (Sivarajah u.a. 2017; Wirth und Hipp 2000) beim Umgang mit *Big Data* beinhaltet die Phasen:

1. *Business understanding* (Verstehen der Anforderungen, Zielsetzungen und Fragestellungen an Big Data auf der Basis domänenspezifischen Wissens)
2. *Data understanding* (Daten sammeln, Qualitätscheck, versteckte Information aufdecken, erste Hypothesen generieren)
3. *Data preparation* (Datenbereinigung und Datenaufbereitung für und Festlegen von informatischen Auswertungsverfahren)

4. ‹Modeling› (Modellannahmen festlegen, Erstellen und iteratives Verfeinern eines Datenmodells aufgrund von Modellannahmen, Entwickeln eines Testdesigns)
5. ‹Evaluation› (der hergeleiteten Modelle, Bewerten der Ergebnisse auf der Basis der Datenqualität, der angewendeten Algorithmen und Auswertungsverfahren, Bewerten der Ergebnisse in Bezug auf die Rechercheziele und Auswertungsinteressen)
6. ‹Deployment› Präsentation der (prädikativen) Ergebnisse für Auftraggeber, Projekt-Review, Schlussfolgerungen für künftiges Handeln, Strategieentwicklung

Die Phasen von CRISP-DM fassen z.T. mehrere Phasen des *Data Life Cycles* zusammen, gehen aber auch über diesen hinaus. Insbesondere in der Phase 1 wird die Notwendigkeit der Kontextualisierung von Fragestellungen auf der Basis von ausserinformatischem Expertenwissen deutlich. In Phase 6 wird die Notwendigkeit der Re-kontextualisierung des Analyseprozesses und der Ableitung von Handlungsstrategien begründet. Für informatische Unterrichtsprojekte zum Thema ‹Big Data› kann daraus abgeleitet werden, dass sie interdisziplinär angelegt sein und eine auf den ausserinformatischen Anwendungskontext bezogene Bewertung der Analyseergebnisse beinhalten sollten. Ein solches methodisches Vorgehen weist hohe Affinität zu dem im Frankfurter Dreieck beschriebenen interdisziplinären Konzept zum Umgang mit Digitalen Medien auf.

Curriculare Konzepte

Mittlerweile gibt es sowohl auf nationaler Ebene in Deutschland (Biehler und Schulte 2018; Grillenberger und Romeike 2018) als auch international eine Reihe curricularer Konzepte zu DSE oder ‹Data Literacy›, die diesen, aus ‹Big Data› erwachsenden Anforderungen an informatische Bildung Rechnung tragen.

Magenheim (2018) hat analysiert, inwieweit die GI-Informatik-Bildungsstandards den Anforderungen informatischer Bildung im Bereich DS genügen und inwieweit Ergänzungen erforderlich sind. Grillenberger und Romeike (2018) haben ein an diesen Informatik-Bildungsstandards

orientiertes, theoretisch begründetes Data-Literacy-Kompetenzmodell mit Prozess- und Inhaltsbereichen vorgeschlagen, das sich speziell auf den Themenbereich ‹Big Data› bezieht.

Die Notwendigkeit, ein interdisziplinäres Curriculum für ‹Data Science Education›, das informatische, mathematische und darüber hinaus relevante gesellschaftliche oder naturwissenschaftliche Gegenstandsbereiche enthält, haben Biehler und Schulte (2018) vorgeschlagen. Als wichtige gesellschaftliche Aspekte, die im Rahmen von ‹Big Data› Projekten thematisiert werden können, und die zugleich zentrale Elemente von Medienkompetenz seien, werden u.a. genannt ‹Kompetenter Umgang mit persönlichen Daten›, ‹Erkennen von fake news›, oder ‹Statistic Literacy›, als Befähigung zur kritischen Beurteilung von journalistischen oder wissenschaftlichen Studien, die sich auf die Analyse von ‹Big Data› stützen.

Einen engen Bezug zwischen dem *Data Life Cycle* und in ‹Data Literacy› / ‹Data Science Education› beschriebenen ‹Schlüsselqualifikationen› (key abilities) haben Ridsdale u.a. (2015). Diese darin implizit enthaltenen Kompetenzerwartungen an die Lernenden gehörten bisher nicht zum Standardrepertoire von nationalen oder internationalen Informatikcurricula.

In dem neuen Informatik-CSTA-K12 Curriculum für die USA werden Kompetenzerwartungen mit explizitem Bezug zu ‹Big Data› angesprochen:

«Use data analysis tools and techniques to identify patterns in data representing complex systems. For example, identify trends in a dataset representing social media interactions, movie reviews, or shopping patterns», «Select data collection tools and techniques to generate data sets that support a claim or communicate information.», «Evaluate the ability of models and simulations to test and support the refinement of hypotheses» (Computer Science Teachers Association 2017).

Weitere Bezüge zu ‹Big Data› finden sich in diesem Curriculum mit den Themenbereichen ‹Collection, Visualization & Transformation› (of Data), ‹Evaluate Hypotheses› und ‹Inference & Models and Algorithms›. Auch in Informatikcurricula für Schulen in Grossbritannien wird die Bedeutung des Themas ‹Big Data› gesehen:

«The opportunities provided by new computing curricula coupled with advances in technologies and analytical tools with which to mine big datasets, and the increasingly interdisciplinary nature of educational research, offer enormous scope for advancing computing teaching and learning» (Royal Society 2017, 96).

In allen diesen Curricula finden sich zudem Hinweise auf die interdisziplinären Bezüge des Themas «Big Data». Es wird vorgeschlagen, das Thema «Big Data» im Unterricht mit «real data» und «real projects» im Rahmen einer «real-world application» umzusetzen (Song und Zhu 2016). Im Rahmen solcher Projekte sollten dann ökonomische, gesellschaftliche und kulturelle Implikationen in die Projekte mit einbezogen werden. Viele Autoren und Autorinnen stellen einen Zusammenhang zwischen «Data Literacy» und «Media Literacy» her, besonders bezüglich des Umgangs mit persönlichen Daten und den Methoden zur Datensicherheit und des Datenmanagements (Yuval Noah Harari 2017; Pangrazio und Selwyn 2019; O'Neil 2017; Aoun 2017).

Tools für multiperspektivische interdisziplinäre Projekte

Unabhängig davon, ob das Thema «Big Data» Gegenstand des Informatikunterrichts ist, in einem fächerübergreifenden Projekt oder integriert in einem Unterrichtsfach realisiert wird, es sind im Sinne des Frankfurter Dreiecks neben den themenbezogenen Aspekten jeweils auch informatische und mediale Aspekte betroffen. Letztere besonders dann, wenn Analysewerkzeuge als digitale Medien und als hybride Informatiksysteme im oben beschriebenen Sinne selbst zum Unterrichtsgegenstand werden. Ihnen kommt dann im Unterricht eine Doppelfunktion zu: Analysetool zur Bearbeitung von themenbezogenen Fragestellungen im Zusammenhang mit grossen Datensammlungen und digitales Medium und Informatiksystem mit Bedienungsfunktionen und innerer Struktur (z.B. Algorithmen, Visualisierung von Daten etc.). Je nachdem, über welche informatischen Kompetenzen die Lerngruppe verfügt, können das sehr einfach zu bedienende Analyse-Tools sein, die in verschiedenen Unterrichtsfächern schon ab der Jahrgangsstufe 5 einsetzbar sind, oder komplexere professionelle

Systeme, die sich eher für die Sekundarstufe II bzw. den universitären Bereich eignen. Einen Überblick über für den Unterricht in der Sekundarstufe geeignete Werkzeuge, die das Sammeln und Aggregieren von Daten sowie deren statistische Auswertung und die Visualisierung von Ergebnissen unterstützen findet sich z.B. in Magenheimer und Schulte (2020).

Mittels dieser Tools lassen sich «Big-Data»-Projekte im Unterricht realisieren, die sich inhaltlich an der Lebenswelt der Schüler und Schülerinnen und wichtigen natur- oder sozialwissenschaftlichen Themen orientieren. Als einige Projektbeispiele seien genannt:

- Analyse von Social Media Daten z.B. von Twitter (Sentiment-Analyse);
- Erstellen von Wahlprognosen auf der Grundlage früher erhobener Daten;
- Auswertung der von Studierenden mit mobilen Geräten erhobenen Daten zu spezifischen wissenschaftlichen Fragestellungen;
- Auswertung von freigegebenen personenbezogenen GPS-, medizinischen und physiologischen Daten, die mit mobilen Geräten erfasst werden;
- Auswertung von Verkehrsflussdaten (z.B. Anwendung der Mustererkennung von Webcams);
- Analyse von Textdokumenten über semantische Ähnlichkeit (u.a. Plagiate);
- Mustererkennung in Bildern und Videos;
- Analyse von Wetterdaten;
- Internationale Projekte zu astronomischen und physikalischen Satellitendaten;
- Prognose von Sportergebnissen auf der Basis früherer Resultate.

Diese Liste von Projektideen liesse sich noch vielfach erweitern. Sie zeigt, dass mit den vorhandenen Analysetools spannender Unterricht zu verschiedenen Anwendungsgebieten von «Big Data» realisierbar ist, und dass dabei neben den fachbezogenen auch mediale und informatische Sichtweisen auf die eingesetzten Tools und den Unterrichtsgegenstand ermöglicht werden.

Literatur

- Albrecht, Steffen, und Christoph Revermann. 2016. «Digitale Medien in der Bildung». 171. TAB-Arbeitsbericht. Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag.
- Alharthi, Abdulkhaliq, Vlad Krotov, und Michael Bowman. 2017. «Addressing Barriers to Big Data». *Business Horizons* 60 (3): 285–92. <https://doi.org/10.1016/j.bushor.2017.01.002>.
- Aoun, Joseph E. 2017. *Robot-proof: higher education in the age of artificial intelligence*. Cambridge, MA: The MIT Press.
- Aßmann, Sandra, Nils Brüggem, Valentin Dander, Harald Gapski, Gerda Sieben, Angela Tillmann, und Isabel Zorn. 2017. «Digitale Datenerhebung und -verwertung als Herausforderung für Medienbildung und Gesellschaft. Ein medienpädagogisches Diskussionspapier zu Big Data und Data Analytics». In *Medienpädagogik*, herausgegeben von Christine Trültzsch-Wijnen, 183–92. Nomos Verlagsgesellschaft mbH & Co. KG. <https://doi.org/10.5771/9783845279718-183>.
- Biehler, Rolf, und Carsten Schulte. 2018. «Perspectives for an interdisciplinary data science curriculum in German Secondary Schools». In *Paderborn Symposium on Data Science Education at School Level 2017: The Collected Extended Abstracts*, herausgegeben von Rolf Biehler, Carsten Schulte, Birte Heinemann, Lea Budde, Susanne Podworny, Daniel Frischemeier, und Thomas Wassong, 2–14. <http://doi.org/10.17619/UNIPB/1-374>.
- Bigdatasoc. 2019. «Call for Special Theme Proposals for Big Data & Society». *Big Data & Society*. 2019. <http://bigdatasoc.blogspot.com/2019/06/call-for-special-theme-proposals-for.html>.
- Brinda, Torsten, Niels Brüggem, Ira Diethelm, Thomas Knaus, Sven Kommer, Christine Kopf, Petra Missomelius, Rainer Leschke, Friederike Tilemann, und Andreas Weich. 2019. «Frankfurt-Dreieck zur Bildung in der digital vernetzten Welt». *Keine Bildung ohne Medien*. 2019. <https://www.keine-bildung-ohne-medien.de/frankfurter-dreieck/>.
- Brinda, Torsten, Ira Diethelm, Rainer Gemulla, Ralf Romeike, Johannes Schöning, und Carsten Schulte. 2016. «Dagstuhl-Erklärung: Bildung in der digitalen vernetzten Welt». *GI – Gesellschaft für Informatik e.V.* 2016. <https://www.gi.de/aktuelles/meldungen/detailansicht/article/dagstuhl-erklaerung-bildung-in-der-digitalen-vernetzten-welt.html>.
- Brüggem, Niels. 2016. «Big Data als Herausforderung für die Medienpädagogik». *tv diskurs, Verantwortung in audiovisuellen Medien*, 20 (78): 18–21.
- Schulte, Carsten, und Lea Budde. 2018. «A Framework for Computing Education: Hybrid Interaction System: The Need for a Bigger Picture in Computing Education». In *Proceedings of the 18th Koli Calling International Conference on Computing Education Research – Koli Calling '18*, 1–10. Koli, Finland: ACM Press. <https://doi.org/10.1145/3279720.3279733>.

- Chen, Min, Shiwen Mao, und Yunhao Liu. 2014. «Big Data: A Survey». *Mobile Networks and Applications* 19 (2): 171–209. <https://doi.org/10.1007/s11036-013-0489-0>.
- Computer Science Teachers Association. 2017. «K12 Computer Science Standards, Revised 2017». gehalten auf der CSTA Annual Conference 2017. <https://www.doe.k12.de.us/cms/lib/DE01922744/Centricity/Domain/176/CSTA%20Computer%20Science%20Standards%20Revised%202017.pdf>.
- Cuquet, Martí, Guillermo Vega-Gorgojo, Hans Lammerant, Rachel Finn, und Umair ul Hassan. 2017. «Societal impacts of big data: challenges and opportunities in Europe». arXiv:1704.03361 [cs], April. <http://arxiv.org/abs/1704.03361>.
- De Mauro, Andrea, Marco Greco, und Michele Grimaldi. 2015. «What is big data? A consensual definition and a review of key research topics». In (*IC-ININFO 2014*): *Proceedings of the 4th International Conference on Integrated Information*, 97–104. Madrid, Spain. <https://doi.org/10.1063/1.4907823>.
- De Veaux, Richard D., Mahesh Agarwal, Maia Averett, Benjamin S. Baumer, Andrew Bray, Thomas C. Bressoud, Lance Bryant, Lei Z. Cheng, Amanda Francis, und Robert Gould. 2017. «Curriculum guidelines for undergraduate programs in data science». *Annual Review of Statistics and Its Application*, Nr. 4: 15–30.
- dpa. 2019. «Fünf Milliarden Dollar: Rekordstrafe für Facebook wegen Cambridge-Analytica-Datenskandal». *Meedia*, Juli.
- Eder, Sabine, Claudia Mikat, und Angela Tillmann, Hrsg. 2017. *Software takes command: Herausforderungen der „Datafizierung“ für die Medienpädagogik in Theorie und Praxis*. Schriften zur Medienpädagogik 53. München: kopaed.
- EU Expertengruppe für künstliche Intelligenz. 2019. «Ethik-Leitlinien für eine vertrauenswürdige KI. Europäische Kommission». https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60425.
- Faundeen, John L. 2012. «Data Life Cycle Models and Concepts CEOS Version 1.1». <https://my.usgs.gov/confluence/download/attachments/82935852/Data%20Lifecycle%20Models%20and%20Concepts%20v11.docx?api=v2>.
- GI – Gesellschaft für Informatik e.V. 2008. «Bildungsstandards Informatik für die Sekundarstufe I: Empfehlungen der Gesellschaft für Informatik e.V.» *LOG IN* 28.
- GI – Gesellschaft für Informatik e.V. 2016. «Empfehlungen für Bachelor- und Masterprogramme im Studienfach Informatik an Hochschulen». *GI-Empfehlungen*. Bonn.
- Grapski, Harald, Thomas Tekster, und Monika Elias. 2018. «Bildung für und über Big Data». Gutachten im Rahmen von ABIDA – Assessing Big Data. Marl: Grimme-Institut.
- Grillenberger, Andreas, und Ralf Romeike. 2018. «Developing a Theoretically Founded Data Literacy Competency Model». In *Proceedings of the 13th Workshop in Primary and Secondary Computing Education on – WiPSCE '18*, 1–10. Potsdam, Germany: ACM Press. <https://doi.org/10.1145/3265757.3265766>.

- Harari, Yuval Noah. 2017. *Homo Deus: A Brief History of Tomorrow*. 1. London: Vintage.
- Heckmann, Dirk. 2019. «Datenverwertung und Datenethik». *Aus Politik und Zeitgeschichte* 69 (24–26).
- Herzig, Bardo. 2016. «Medienbildung und Informatische Bildung – Interdisziplinäre Spurensuche». *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung* 25 (Computer Science Education), 59–79. <https://doi.org/10.21240/mpaed/25/2016.10.28.X>.
- Klafki, Wolfgang. 1993. «Allgemeinbildung heute». *Pädagogische Welt*, 47:28–33.
- Kolmar, Martin, und Johannes Binswanger. 2019. «Algorithmen: Sie entscheiden über unsere Leben». *Zeit Online*. 2019. <https://www.zeit.de/wirtschaft/2019-07/algorithmus-facebook-google-datensicherheit>.
- Lepri, Bruno, Jacopo Staiano, David Sangokoya, Emmanuel Letouzé, und Nuria Oliver. 2017. «The Tyranny of Data? The Bright and Dark Sides of Data-Driven Decision-Making for Social Good». In *Transparent Data Mining for Big and Small Data*, herausgegeben von Tania Cerquitelli, Daniele Quercia, und Frank Pasquale, 32:3–24. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-54024-5_1.
- Magenheim, Johannes. 2008. «Systemorientierte Didaktik der Informatik Sozio-technische Informatiksysteme als Unterrichtsgegenstand?» Informatische Ideen im Mathematikunterricht: Bericht über die 23. Arbeitstagung des Arbeitskreises «Mathematikunterricht und Informatik» in der Gesellschaft für Didaktik der Mathematik e.V. vom 23. bis 25. September 2005 in Dillingen an der Donau. Hildesheim: Franzbecker Verlag.
- Magenheim, Johannes. 2018. «Data science as a school subject in secondary education from the perspective of computer science education». In *Paderborn Symposium on Data Science Education at School Level 2017: The Collected Extended Abstracts*, herausgegeben von Rolf Biehler, Lea Budde, Daniel Frischemeier, Birte Heinemann, und Susanne Podworny, 95–102.
- Magenheim, Johannes, und Carsten Schulte. 2020. «Data Science Education». In *Encyclopedia of Education and Information Technologies*, herausgegeben von Arthur Tatnall. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-319-60013-0>.
- Mantelero, Alessandro. 2018. «AI and Big Data: A Blueprint for a Human Rights, Social and Ethical Impact Assessment». *Computer Law & Security Review* 34 (4): 754–72. <https://doi.org/10.1016/j.clsr.2018.05.017>.
- Mau, Steffen. 2018. *Das metrische Wir: über die Quantifizierung des Sozialen*. Bonn: Bundeszentrale für politische Bildung.
- Mittelstadt, Brent Daniel, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, und Luciano Floridi. 2016. «The Ethics of Algorithms: Mapping the Debate». *Big Data & Society* 3 (2). <https://doi.org/10.1177/2053951716679679>.

- O'Neil, Cathy. 2017. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. London: Penguin Books.
- Pangrazio, Luci, und Neil Selwyn. 2019. «'Personal data literacies': A critical literacies approach to enhancing understandings of personal digital data». *new media & society* 21 (2): 419–37. <https://doi.org/10.1177/1461444818799523>.
- Ridsdale, Chantel, James Rothwell, Mike Smit, Hossam Ali-Hassan, Michael Blieme, Dean Irvine, Daniel Kelley, Stan Matwin, und Brad Wuetherick. 2015. «Strategies and best practices for data literacy education: knowledge synthesis report». Dalhousie University. <http://hdl.handle.net/10222/64578>.
- Riss, Uwe V, Johannes Magenheimer, Wolfgang Reinhardt, Tobias Nelkner, und Knut Hinkelmann. 2011. «Added Value of Sociofact Analysis for Business Agility». In *Papers from the AAAI 2011 Spring Symposium*. Bd. SS-11-03: 13.
- Ropohl, Günter. 2009. *Allgemeine Technologie. Eine Systemtheorie der Technik*. Karlsruhe: Universitätsverlag Karlsruhe.
- Royal Society (Great Britain). 2017. *After the Reboot: Computing Education in UK Schools*. <https://royalsociety.org/~media/policy/projects/computing-education/computing-education-report.pdf>.
- Schiersmann, Christiane, Johannes Busse, und Detlef Krause. 2002. «Medienkompetenz – Kompetenz für Neue Medien Studie und Workshop». 12. Arbeitsstab Forum Bildung in der Geschäftsstelle der Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung. <http://www.blk-bonn.de/papers/forum-bildung/band12.pdf>.
- Schulte, Carsten, Johannes Magenheimer, Kathrin Müller, und Lea Budde. 2017. «The design and exploration cycle as research and development framework in computing education». In *2017 IEEE Global Engineering Education Conference (EDUCON)*, 867–76. Athens, Greece: IEEE. <https://doi.org/10.1109/EDUCON.2017.7942950>.
- Sivarajah, Uthayasankar, Muhammad Mustafa Kamal, Zahir Irani, und Vishanth Weerakkody. 2017. «Critical Analysis of Big Data Challenges and Analytical Methods». *Journal of Business Research* 70: 263–86. <https://doi.org/10.1016/j.jbusres.2016.08.001>.
- Song, Il-Yeol, und Yongjun Zhu. 2016. «Big Data and Data Science: What Should We Teach?». *Expert Systems* 33 (4): 364–373. <https://doi.org/10.1111/exsy.12130>.
- Tulodziecki, Gerhard. 2016a. «Aktuelle Debatten beim GMK-Forum 2015 im „Rückspiegel“ Welchen Lösungsbeitrag können medienpädagogische Grundlagen leisten?». In *Kommunikationskulturen in digitalen Welten*, herausgegeben von Marion Brüggemann, Dorothee M. Meister, und Thomas Knaus. Bd. 52. Schriften zur Medienpädagogik. München: kopaed.
- Tulodziecki, Gerhard. 2016b. «Konkurrenz oder Kooperation? Zur Entwicklung des Verhältnisses von Medienbildung und informatischer Bildung». *Medienpädagogik: Zeitschrift für Theorie und Praxis der Medienbildung* 25 (Computer Science Education), 7-25. <https://doi.org/10.21240/mpaed/25/2016.10.25.X>.

- Wirth, Rüdiger, und Jochen Hipp. 2000. «CRISP-DM: Towards a standard process model for data mining». In *Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining*, 29–39.
- Zhang, Yaoxue, Ju Ren, Jiagang Liu, Chugui Xu, Hui Guo, und Yaping Liu. 2017. «A Survey on Emerging Computing Paradigms for Big Data». *Chinese Journal of Electronics* 26 (1): 1–12. <https://doi.org/10.1049/cje.2016.11.016>.
- Zorn, Isabel. 2011. «Zur Notwendigkeit der Bestimmung einer auf Digitale Medien fokussierten Medienkompetenz und Medienbildung». *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung* 20 (Medienbildung – Medienkompetenz), 175-209. <https://doi.org/10.21240/mpaed/20/2011.09.19.X>.