

Zur (unterschätzten) Eigendynamik von Forschungsdateninfrastrukturen

Ein Kommentar zu den „Empfehlungen zur Archivierung, Bereitstellung und Nachnutzung von Forschungsdaten im Kontext erziehungs- und bildungswissenschaftlicher sowie fachdidaktischer Forschung“

*Sigrid Hartong, Claudia Machold & Patricia Stošić
für den Vorstand der Sektion 3 – Interkulturelle und
International Vergleichende Erziehungswissenschaft*

Einleitung – wichtige Schritte, bleibender Diskussionsbedarf

Am 11. März 2020 wurden die „Empfehlungen zur Archivierung, Bereitstellung und Nachnutzung von Forschungsdaten im Kontext erziehungs- und bildungswissenschaftlicher sowie fachdidaktischer Forschung“ im Sinne einer gemeinsamen Stellungnahme der Deutschen Gesellschaft für Erziehungswissenschaft (DGfE), der Gesellschaft für Empirische Bildungsforschung (GEBF) sowie der Gesellschaft für Fachdidaktik (GFD) veröffentlicht. Die Empfehlungen sind Ergebnis einer Aufforderung der Deutschen Forschungsgemeinschaft (DFG) an die Fachgesellschaften, Leitlinien zum Forschungsdatenmanagement zu entwickeln. Entsprechend bedeutsam ist das nun entstandene Empfehlungspapier sowohl für die Forschungsförderung als auch allgemein für Wissenschaft und Wissenschaftspolitik.

Zunächst verdient die Diskussions- und Integrationsarbeit teils signifikant unterschiedlicher Forschungstraditionen, die im Papier deutlich zum Ausdruck kommt, unbedingte Anerkennung. So sprechen sich die Autorinnen und Autoren des Papiers zwar ganz klar für eine Stärkung der Archivierung, des Verfügbarmachens sowie der Nachnutzbarkeit von Forschungsdaten aus, verfolgen aber ebenso den Anspruch, Risiken bzw. besondere Herausforderungen ausreichend in den Blick zu nehmen. Hierbei nennen sie Spezifika erziehungs- und bildungswissenschaftlicher sowie fachdidaktischer Forschung, die etwa eine „[...] besondere Reflexion der datenschutzrechtlichen Bestimmungen“ sowie „forschungsethischer Implikationen“ (S. 3) erforderlich machten, oder Herausforderungen der gleichberechtigten Wertschätzung von Primär- und Sekundärforschung (ebd.). Gleichzeitig betonen die Autorinnen und Autoren an mehreren Stellen explizit, dass die Empfehlungen keinesfalls zu einem „Zwang zur Nachnutzbarkeit“ (S. 9) von Daten bzw. zu einer möglicherweise

nachteiligen Auslegung im Rahmen von Vergabe- und Begutachtungsverfahren (vgl. S. 11) führen, das *Soll* also nicht in ein *Muss* umschlagen dürfe.

Die Berücksichtigung all dieser (und weiterer im Papier genannter) Punkte bildet aus unserer Perspektive die Grundlage für eine fruchtbare Auseinandersetzung und bewusste Gestaltung von Fragen des Forschungsdatenmanagements. Darüber hinaus ist es jedoch notwendig, so unser Anliegen, noch wesentlich stärker die (strukturellen wie normativen) Eigendynamiken in den Blick zu nehmen, die aus unserer Sicht mit einer Etablierung und Ausdifferenzierung von Dateninfrastrukturen in Gang gesetzt werden (können). An diesem Punkt möchten wir im Folgenden mit einer kritischen Kommentierung der veröffentlichten Empfehlung ansetzen, indem wir zunächst gegenwärtige Entwicklungen in den Blick nehmen, die Dateninfrastrukturen als zunehmend einflussreiche Elemente von Wissenschaftssteuerung und -praxis etablieren (Abschnitt 2). Vor diesem Hintergrund heben wir im nächsten Schritt Aspekte des Empfehlungspapiers hervor, die insofern weiter problematisiert werden müssen, als sie Gefahr laufen, dieser zunehmenden Einflussnahme mit ihren nicht-intendierten Effekten in die Hände zu spielen (Abschnitt 3). Abschließend formulieren wir sodann Positionen und Maßnahmen, die Teil einer notwendig weiterhin zu führenden und zu intensivierenden Debatte um Forschungsdatenmanagement sein können (Abschnitt 4).

Dateninfrastrukturen als zunehmend einflussreiche Elemente von Wissenschaftssteuerung und -praxis

Während die Produktion und Auswertung von Daten seit jeher Gegenstand empirischer Forschung gewesen ist, so lässt sich in den letzten Jahren ein gravierender Wandlungsprozess identifizieren. Hierbei ist neben einer wachsenden Transferorientierung im Sinne der verbesserten Nutzbarmachung wissenschaftlicher Erkenntnisse für *evidenzbasierte* Politik und Praxis (vgl. auch Jorntz 2008, S. 206) die Schaffung sogenannter „Informations- und Forschungsdateninfrastrukturen“ (BMBF 2018:23) zunehmend ins Zentrum der politischen Aufmerksamkeit und Finanzierung gerückt. So stellt etwa die *Verbesserung von Forschungsinfrastrukturen und Forschungsdaten* (vgl. BMBF 2018, S. 23) ein eigenständiges Ziel des Rahmenprogramms empirische Bildungsforschung des Bundesministeriums für Bildung und Forschung (BMBF) dar, in dessen Kontext seit 2013 insbesondere der Verbund Forschungsdaten Bildung massiv ausgebaut wurde.¹

Den aufgeführten Förderprogrammen und Debatten unterliegt meist ein funktionales, oftmals technisch-neutral gerahmtes Verständnis von Dateninfrastrukturen, welches sich in entsprechenden Argumentationen ausdrückt:

1 <https://www.forschungsdaten-bildung.de> [Zugriff: 20. Oktober 2020]

Daten seien ‚unter einem Dach‘ verfügbar, Effizienz, Nutzungsfreundlichkeit oder Qualitätssicherung könnten durch standardisierte, rechtlich abgesicherte Verfahren (siehe auch BMBF 2020) gesteigert werden und dergleichen. Die Daten selbst erscheinen dann als objektiviertes Rohmaterial, das – einmal erhoben und gut aufbereitet bzw. dokumentiert – im Prinzip beliebig oft für weitergehende Analysen genutzt werden kann.

Dies steht einer Perspektive auf Dateninfrastrukturen entgegen, die sich insbesondere in der kritischen Datafizierungs- und Digitalisierungsforschung der letzten Jahre entwickelt hat (z. B. Sellar 2015). Hierbei werden zwar ebenso die organisationalen Prozesse von Datenproduktion, -verarbeitung oder -nutzung in den Fokus gerückt, jedoch werden diese Prozesse nicht als neutral-technologische Instrumente der Wissensproduktion, -organisation und -optimierung betrachtet, sondern als wertbehaftete, in ihrer Rationalität und Normativität machtvolle Apparate der Selektion und In-Formation von (Bildungs- bzw. Wissenschafts-)Welt (vgl. auch Ruppert et al. 2017). Folgt man dieser Argumentation, geht es weniger um die Daten selbst (als vermeintliche Repräsentation der Bildungswelt), sondern vielmehr um die komplexen soziotechnischen De- und Rekontextualisierungsprozesse, die im Zusammenspiel von Objekten (die Daten selbst, Hard- und Software, aber auch Politikfragmente wie z. B. Richtlinien, Standards etc.), Subjekten (Technikerinnen und Techniker, Verwalterinnen und Verwalter, Wissenschaftlerinnen und Wissenschaftler, Vermittlungsakteurinnen und Vermittlungsakteure etc.) sowie einer Reihe von Datenpraktiken (Modellierung, Datenaufbereitung etc.) auch nicht intendierte Nebeneffekte und eine oftmals unterschätzte Wirkmächtigkeit erzeugen (vgl. auch Hartong 2018, S. 135). Diese Wirkmächtigkeit ist zum einen *onto-politisch*, d. h. sie erzeugt eine bestimmte Sicht bzw. Wahrnehmbarkeit von (guten) erziehungswissenschaftlichen Daten, (guter) Forschung, ihren Gegenständen etc. (Scheel et al. 2019, S. 681). Zum anderen ist sie *performativ*, d. h. jede Infrastruktur macht bestimmte Dinge (nicht) möglich, etwa was die Vorstrukturierung wissenschaftlicher oder politischer Entscheidungen angeht (vgl. z. B. Gorur/Dey 2020, S. 3f.; van de Oudeweetering/Decuypere 2019, S. 5).

Unter einer solchen Perspektive erscheinen sowohl die kontinuierliche Expansion als auch die politisch angestrebte Verdichtung von (Forschungs-)Dateninfrastrukturen zumindest fragwürdig. Denn wo es einerseits gelingen soll, eine wachsende Anzahl ganz unterschiedlicher Daten – von *Big Data* bis hin zu ethnographischen Kontextdaten – zu integrieren, bedarf es für eben diese Integration enormer Anstrengungen der Zentralisierung, Koordination und Standardisierung, die wiederum die (mögliche) Vielfalt, Bildung und Erziehung datentechnisch zu erfassen und damit *in Wert zu setzen*, systematisch reduzieren und gleichzeitig (entsprechend ihrer Entstehungslogik) de- bzw. (entsprechend der Infrastrukturlogik) rekontextualisieren.

Beispiele für diese reduktive Kraft der Standardisierung sind Daten-Interoperabilitätsrahmen oder die Festlegung von Meta-Daten (vgl. Kubicek et al.

2019) ebenso wie der gesamte Bereich des Datenmanagements. Hierbei wird zunehmend kleinschrittig festgelegt, welche Daten in welchen Formaten überhaupt in die Infrastruktur aufgenommen bzw. verarbeitet werden können und welche Datenpraktiken (etwa *data cleaning*, Ratner/Ruppert 2019) hierfür zu welchem Zeitpunkt erforderlich sind. Derartige Datenpraktiken spielen entsprechend im Forschungsprozess eine zunehmend umfangreiche Rolle. Als Anschauungsbeispiel soll hier der Umgang mit offenen Antworten dienen, wie ihn der Verbund Forschungsdaten Bildung empfiehlt:

„Offene Fragen (sog. String-Variablen) sind Fragen mit offener Antwortmöglichkeit. Sie bilden bezüglich der Codierung unter Umständen eine Ausnahme, da sie sich teilweise nicht in einfacher und sinnvoller Weise in numerische Codes umwandeln lassen. Sie werden dann u. U. nicht umcodiert, um keinen Informationsverlust zu verursachen, sondern in ursprünglicher Form belassen, dann aber ohne statistische/quantitative Auswertungsmöglichkeiten, nur teilweise umcodiert oder aber vollständig aus dem Datensatz entfernt. Zudem bergen sie das Risiko, dass sie Informationen enthalten, die zu einer Identifizierung der Teilnehmer/-innen führen könnten. Offene Angaben sind daher sorgfältig auf datenschutzrechtliche Probleme zu untersuchen.“ (Trixa/Ebel 2015, S. 7)

Mit Gorur und Dey (2020, S. 5) ließe sich also überspitzt formulieren: „[...] [I]t is not just the system that becomes increasingly ‚user friendly‘, but the user is also made friendly toward the system.“ Eine zunehmend zentrale Rolle spielen hierbei wiederum Akteurinnen und Akteure, die als *neue Knotenpunkte der Datensammlung* fungieren bzw. definieren, für wen welche Daten unter welchen Voraussetzungen zugänglich gemacht, welche Standards gesetzt und wie diese erfolgreich an potenzielle Nutzerinnen und Nutzer vermittelt werden (siehe bspw. Verbund Forschungsdaten Bildung 2020).

So zentral all diese Prozesse für eine ‚reibungslose‘, möglichst ‚optimierte‘ Funktionsweise von Dateninfrastrukturen sind, so gewinnen sie dabei doch auch und zwar in zunehmendem Maße Einfluss darauf, wie Forschungsgegenstände, -fragen und -prozesse letztendlich zu sehen sind und welches Leitbild *guter Forschung, guter Forschender* bzw. guter Datenmanagerinnen und Datenmanager (und zwar auch in Politik und Wissenschaftsförderung) vorstrukturiert bzw. verbreitet wird. Anders formuliert: Wenn sich Datenformate oder -praktiken kaum mehr widersprechen, werden sie als solche mitunter auch weniger hinterfragt, sondern entfalten sukzessive Mechanismen der Normalisierung.

All diese Überlegungen sollen verdeutlichen, dass es bei einer kritischen Auseinandersetzung mit Perspektiven auf Datenintegration und Wiederverwendung eben nicht nur zu Herausforderungen bezüglich des Datenschutzes oder der Forschungsethik kommt. Im Gegenteil können die aktuell oftmals gewählten Strategien bei der Beantwortung von Datenschutz- und Forschungsethikfragen – nämlich ein nochmals stärkerer Fokus auf Standardisierung, auf die Produktion forschungsethischer ‚Checkboxes‘ oder die Entwicklung von Anonymisierungskonzepten, die eine Weitergabe von ‚sicheren‘ Daten erlau-

ben – durchaus zu einer weiteren Verstärkung oben genannter Dynamiken führen. Weiter kann es aber auch sein, dass es durch die aufwendigere Verarbeitung qualitativer, insbesondere ethnographisch geprägter Daten zu einer geringeren Repräsentanz derartiger Daten innerhalb der Forschungsdatenzentren kommt, oder aber dazu, dass Forschende von der Erhebung derartiger Daten zunehmend Abstand nehmen.

Perspektiven des Empfehlungspapiers auf Dateninfrastrukturen

Vor dem Hintergrund der im letzten Abschnitt dargelegten Perspektive auf mögliche Eigendynamiken und normative Implikationen von Dateninfrastrukturen sehen wir eine Notwendigkeit, die „Empfehlungen zur Archivierung, Bereitstellung und Nachnutzung von Forschungsdaten im Kontext erziehungs- und bildungswissenschaftlicher sowie fachdidaktischer Forschung“ systematisch weiter zu diskutieren und nachzubearbeiten.

So findet sich *erstens* auch im Empfehlungspapier ein Verständnis von „unbearbeiteten Rohdaten“ (S. 5), welches die soziale Bedingtheit bzw. *Gemachtheit* von Daten eher ausblendet. Als besonders beachtenswert erscheint uns hierbei, dass neben Mess- oder Log-Daten sogar Beobachtungsprotokolle oder Feldnotizen unter dieser Kategorie gefasst werden.

Zweitens werden auch im Empfehlungspapier problematische Aspekte von Dateninfrastrukturen tendenziell auf Fragen von Datenschutz und Forschungsethik reduziert (z. B. S. 3), wenngleich das Papier dezidiert argumentiert, dass sich diese Probleme gerade in der erziehungswissenschaftlichen Forschung eben nicht einfach durch Anpassungen der Datenverarbeitung lösen lassen. Nichtsdestotrotz sehen wir auch hier die Gefahr, dass sich die Beachtung von Datenschutz als ‚problemfreie‘ Datennutzung interpretieren ließe. Aus unserer Sicht wäre entsprechend wichtig, auch Datenschutzrichtlinien bzw. ethische Regularien (z. B. das Thema Pseudonymisierung oder die Zustimmung durch Interviewte) als wissenschaftlich und politisch bedingt zu verstehen, die also auch als Gegenstände selbst immer wieder kritisch hinterfragt werden müssen (anstatt sie zur ‚Checkbox‘ zu machen).

Wie eingangs bereits erwähnt, spricht sich das Empfehlungspapier *drittens* ganz klar dafür aus, dass „[...] die Nachnutzung von Forschungsdaten nicht als generelle Forderung an Forschungsprojekte postuliert werden [kann]. Es muss die Möglichkeit bestehen bleiben, Forschungsdaten nicht bzw. nur eingeschränkt zur Verfügung zu stellen“ (S. 17). An dieser Stelle geht das Papier allerdings zum einen nicht über eine Sollempfehlung hinaus und läuft damit Gefahr, die oben skizzierten Dynamiken auszublenden. Zum anderen durchziehen auch die Empfehlungen selbst normative Implikationen, die sich etwa in den angebotenen Checklisten, aber ebenso in gewählten Begrifflichkeiten finden – beispielsweise im verwendeten Begriff des Nachnutzungs*potenzials*.

Mögliche Datennutzungsvarianten, bei denen keine Daten an Forschungsdatenzentren übergeben werden (können oder sollen), erwecken entsprechend zumindest indirekt den Eindruck, als verfüge ein solches Projekt eben über kein Nachnutzungspotenzial bzw. „keine Nachnutzungsmöglichkeiten“ (S. 18). Ein derartiger Eindruck spiegelt sich gleichermaßen in der Abbildung zum Ablauf des Datenbereitstellungsprozesses (S. 18) wider, in welcher die Entscheidung für eine Nicht-Übergabe von Daten als eine Art abgespaltener Sonderfall erscheint. Deutlich klarer könnte hier gemacht werden, dass Daten stattdessen vielseitig nachgenutzt werden (können), auch ohne sie an Forschungsdatenzentren zu übergeben, sowie dass eine Entscheidung für oder gegen eine Weitergabe sich nicht nur aus „Nachteilen für Beforschte/Forschende sowie Primärforschungsprojekte“ (Abb. S. 18) begründen muss.

Eine ähnliche Spannung normativer Positionierung des Empfehlungspapiers findet sich in der Frage, wie mit qualitativem Wissen von Forschenden umgegangen werden soll, das einen zentralen Aspekt erziehungswissenschaftlicher Forschung markiert. Während auf der einen Seite die alleinige Thematisierung und Anerkennung dieser Art Wissen im Papier einen wichtigen Schritt darstellt und gleichzeitig eine gewinnbringende Möglichkeit zur gezielten Auseinandersetzung mit Chancen und Grenzen von Dateninfrastrukturen böte, so enttäuscht die vorgeschlagene Lösung dieser Auseinandersetzung – nämlich die Ergänzung von Datenlieferungen um „Objektdaten“ bzw. erfahrungsbasiertes „Insiderwissen“ (S. 6).

Viertens wird im Papier kaum diskutiert, welche (auch negativen) Folgen sich ggf. ergeben können, wenn sowohl Nutzungsrechte für Daten als auch die Entscheidung über die Passung von Primär- und Sekundärforschung (vgl. S. 19) – was Voraussetzung sein soll für die Datenweitergabe an anfragende Sekundärforschende – an Forschungsdatenzentren externalisiert und innerhalb externer Strukturen wiederum zunehmend zentralisiert, expandiert und mitunter ökonomisch nutzbar gemacht werden. Hier gilt es unseres Erachtens kritisch über Definitionsmacht und Verantwortlichkeiten zu reflektieren. Auch sollte stärker darüber nachgedacht werden, welche Potenziale eines direkten, dezentralen – und damit vielleicht nicht unbedingt effizienten, aber auf anderem Wege gewinnbringenden – Austauschs zwischen Primär- und Sekundärforschenden über Daten, Datenproduktion und Forschungskontext über die Auslagerung ggf. auch ‚verschenkt‘ bzw. zu wenig genutzt werden. So könnte eine Alternative sein, Forschungsdatenzentren stärker als Vernetzungsplattformen zu begreifen, die eben diesen Austausch fördern. Basis könnte auch hier die Hinterlegung von (Daten-)Beschreibungen laufender oder vergangener Forschungsprojekte auf einer Onlineplattform sein, welche interessierte Forscherinnen und Forscher durchsuchen, sich ggf. durch die FDZ beraten lassen und bei Interesse mit den Primärforschenden Kontakt aufnehmen könnten. Auch hier könnte das Zurverfügungstellen von Daten ein Ziel sein, aber ebenso das gezielte Vernetzen der „Menschen hinter den Projekten“.

Gerade weil die Empfehlungen gezielt darauf ausgerichtet sind, Anhaltspunkte für forschungsfördernde Institutionen, insbesondere die Deutsche Forschungsgemeinschaft (DFG) zu bieten, sehen wir das Risiko, dass die normativen Implikationen langfristig zu Lasten der angestrebten Balance und Offenheit und damit der Heterogenität von Forschung (Helsper 2017) gehen. Dies betraf sowohl die Auflage von Forschungsprogrammen und deren Evaluation als auch Begutachtungsprozesse sowie Bewilligungskriterien und damit letztlich die Verteilung von Forschungsgeldern – insgesamt Praktiken, bei denen es zunehmend selbst um kriterienbasierte Entscheidungen geht. In der Tat sind Auswirkungen derartiger Fördernormen, etwa die wachsende Fokussierung auf bzw. kontinuierliche Expansion von Datenmanagementaufgaben, etwa bei BMBF-Forschungsförderungsprogrammen bereits heute deutlich sichtbar.

Ausblick: Anregungen für die notwendige Weiterführung und Intensivierung der Debatte

Hieran anschließend, müsste es daher aus unserer Perspektive darum gehen, die im Empfehlungspapier angestrebte Offenheit durch gezielte Positionierungen und Maßnahmen zu flankieren – sowohl im Rahmen der weitergehenden Standpunktentwicklung in der DGfE, der Debatte in der Fachcommunity sowie im Austausch mit Förderinstanzen und Politik. Derartige Positionierungen und Maßnahmen könnten konkret umfassen:

1. Den bewussten und unbedingten Schutz von *Räumen der Nichtstandardisierung* bzw. von Datenvielfalt als notwendigem Element (nicht nur) erziehungswissenschaftlicher Forschung. Hierzu zählen das bewusste Aufzeigen und Anerkennen gewinnbringender Varianten der Eigenarchivierung von Daten ebenso wie der (zahlreichen) Möglichkeiten (und Notwendigkeiten), Kolleginnen und Kollegen, Bildungspolitik oder -praxis die eigenen Datenpraktiken bzw. auch das eigene Kontextwissen transparent zu machen. Aber es zählt auch dazu, sich kontinuierlich über Sinn, Chancen und Grenzen von Datenmanagement auseinanderzusetzen. Entsprechende Debatten sollten auch in Regelwerken guter Wissenschaft reflektiert werden.
2. Die Sicherstellung entsprechender Offenheit im Kontext von Forschungsförderung, etwa im Kontext der Entwicklung von Leitlinien für Begutachtende.
3. Die weitergehende Problematisierung und rechtliche Regulierung von Datenschutz und Ethik. Gleichzeitig darf die Debatte nicht darin enden, Datenschutz und ethische Fragen zu ‚abhakbaren‘ Bausteinen von Forschung zu machen.
4. Eine weitergehende Reflexion darüber, in welchen Rahmen sich die Auslagerung von Datenverantwortung sowie Entscheidungskompetenz über Datennachnutzung an Forschungsdatenzentren bewegen kann und soll. Hierzu zählen beispielsweise Grenzen der ‚Dateninfrastrukturalisierung‘ von Kon-

- textwissen, aber ebenso Möglichkeiten und Chancen des direkten Austauschs zwischen Forschenden (mit ggf. FDZ als Vermittlungsplattformen).
5. Und schließlich die Förderung dessen, was Gray et al. (2018, S. 1) als *Dateninfrastrukturkompetenz* (*data infrastructure literacy*) beschreiben und von einer reinen Datenkompetenz (*data literacy*) im Sinne eines reflektierten Arbeitens mit Datensätzen abgrenzen. Vielmehr geht es ihnen um die Fähigkeit, die breiteren soziotechnischen Infrastrukturen, durch die Daten in wachsendem Maße erzeugt, gespeichert und analysiert werden, besser zu verstehen, sich mit ihnen – und mit ihrem politischen Charakter – kritisch auseinanderzusetzen, auch um letztendlich aktiv und demokratisch an ihrer Gestaltung mitwirken zu können.

Sigrid Hartong, Prof. Dr., ist Professorin für Soziologie, insb. Transformation von Governance in Bildung und Gesellschaft, an der Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg.

Claudia Machold, Prof. Dr., ist Universitätsprofessorin für Erziehungswissenschaft mit dem Schwerpunkt Kindheitsforschung an der Bergischen Universität Wuppertal.

Patricia Stošić, Dr., ist Wissenschaftliche Mitarbeiterin am Institut für Allgemeine Erziehungswissenschaft, Arbeitsbereich Erziehung, Politik und Gesellschaft an der Goethe-Universität Frankfurt.

Literatur

- BMBF (2018): Rahmenprogramm empirische Bildungsforschung. https://www.bmbf.de/upload_filestore/pub/Rahmenprogramm_empirische_Bildungsforschung.pdf [Zugriff: 8. Juli 2020].
- BMBF (2020): Forschungsdaten Bildung. <https://www.empirische-bildungsforschung-bmbf.de/de/3034.php> [Zugriff: 8. Juli 2020].
- Gorur, Radhika/Dey, Joyeeta (2020): Making the user friendly: the ontological politics of digital data platforms. In: *Critical Studies in Education*, S. 1-15. <https://doi.org/10.1080/17508487.2020.1727544>.
- Hartong, Sigrid (2018): Towards a topological re-assemblage of education policy? Observing the implementation of performance data infrastructures and „centers of calculation“ in Germany. In: *Globalisation, Societies and Education* 16, 1, S. 134-150. <https://doi.org/10.1080/14767724.2017.1390665>.
- Helsper, Werner (2017): Wird die Pluralität in der Erziehungswissenschaft aufgekündigt? In: Baumert Jürgen/Tillmann Klaus-Jürgen (Hrsg.): *Empirische Bildungsforschung (= Zeitschrift für Erziehungswissenschaft, Sonderheft 31)*. Wiesbaden: Springer VS, S. 89-105, https://doi.org/10.1007/978-3-658-13785-4_6.

- Jornitz, Sieglinde (2008): Was bedeutet eigentlich „evidenzbasierte Bildungsforschung“? Über den Versuch, Wissenschaft für Praxis verfügbar zu machen am Beispiel der Review-Erstellung. In: *Die Deutsche Schule* 100, 2, S. 206-216.
- Kubicek, Herbert/Breiter, Andreas/Jarke, Juliane (2019): Daten, Metadaten, Interoperabilität. In: Klenk Tanja/Nullmeier Frank/Wewer Göttrik (Hrsg.): *Handbuch Digitalisierung in Staat und Verwaltung*. Wiesbaden: Springer VS, S. 1-13.
- Oudeweetering, Karmijn van de/Decuypere, Mathias (2019): Understanding openness through (in) visible platform boundaries: A topological study on MOOCs as multiplexes of spaces and times. In: *International Journal of Educational Technology in Higher Education* 16, 1, S. 1-30.
- Ratner, Helene/Ruppert, Evelyn (2019): Producing and projecting data: Aesthetic practices of government data portals. In: *Big Data & Society* 6, 2, S. 2.
- Scheel, Stephan/Ruppert, Evelyn/Ustek-Spilda, Funda (2019): Enacting migration through data practices. In: *Society and Space* 37, 4, S. 579-588. <https://journals.sagepub.com/doi/full/10.1177/0263775819865791>.
- Sellar, Sam (2015): Data infrastructure: a review of expanding accountability systems and large-scale assessments in education. In: *Discourse: Studies in the Cultural Politics of Education* 36, 5, S. 765-777. <https://doi.org/10.1080/01596306.2014.931117>.
- Trixa, Jessica/Ebel, Thomas (2015): Hinweise zur Aufbereitung quantitativer Daten. www.forschungsdaten-bildung.de/get_files.php?action=get_file&file=fdb-informiert-nr-4.pdf [Zugriff: 17. August 2020].
- Verbund Forschungsdaten Bildung (2020): VerbundFDB – Aufbau und Gestaltung einer Forschungsdateninfrastruktur für die empirische Bildungsforschung. www.forschungsdaten-bildung.de/projekt [Zugriff: 8. Juli 2020].