

Datenmanagement

Vera Husfeldt*, David Schiller und Sharon Alt

Das erweiterte Potenzial von Bildungsdaten

<https://doi.org/10.1515/iwp-2020-2116>

Zusammenfassung: Je mehr Bildungsdaten wir sinnvoll kombinieren und verwenden können, desto mehr können wir über das Lernen und über die Qualität unseres Bildungssystems aussagen. Eine interdisziplinäre Forschungsgruppe der Fachhochschule Graubünden (FHGR) hat sich im Rahmen ihres Projekts „Virtual Educational Observatory“ (VEO) zum Ziel gesetzt, das Potenzial von Bildungsdaten durch die Verknüpfung und Weiterverarbeitung vorhandener und zukünftig anfallender Quellen zu erweitern. Der Beitrag beschreibt die Voraussetzungen und Ziele des soeben begonnenen Projekts und geht dabei auch auf die Einbettung in die aktuelle Bildungslage ein, die durch die Corona-Epidemie entstanden ist.

Deskriptoren: Bildungsdaten, Bildungsforschung, Bildungswesen, Daten, Data Warehouse, Datenbasis, Datenschutz, Schule, Schweiz

The expanded potential of educational data

Abstract: The more educational data we can combine and use meaningfully, the more we can say about learning and about the quality of our education system. Within the framework of its “Virtual Educational Observatory” (VEO) project, an interdisciplinary research group at the University of Applied Sciences of the Grisons (FHGR) has set itself the goal of expanding the potential of educational data by linking and processing existing and future sources.

Descriptors: Educational data, Educational research, Education, Data, Data warehouse, Database, Data protection, School, Switzerland

*Kontaktperson: Dr. Vera Husfeldt, Fachhochschule Graubünden, Schweizerisches Institut für Informationswissenschaft, Pulvermühlestrasse 57, 7000 Chur, Schweiz, E-Mail: vera.husfeldt@fhgr.ch

David Schiller, M.A., Fachhochschule Graubünden, Schweizerisches Institut für Informationswissenschaft, Pulvermühlestrasse 57, 7000 Chur, Schweiz, E-Mail: david.schiller@fhgr.ch

Sharon Alt, MSc. BA., Fachhochschule Graubünden, Schweizerisches Institut für Informationswissenschaft, Pulvermühlestrasse 57, 7000 Chur, Schweiz, E-Mail: sharon.alt@fhgr.ch

Élargir le potentiel des données sur l'éducation

Résumé: Si nous sommes en mesure de combiner et d'utiliser un nombre plus important de données sur l'éducation, nous pouvons aussi augmenter les déclarations fondées au sujet de l'apprentissage et de la qualité de notre système éducatif. Dans le cadre de son projet « Virtual Educational Observatory » (VEO), un groupe de recherche interdisciplinaire de la Haute école spécialisée des Grisons (Fachhochschule Graubünden, FHGR) s'est fixé comme objectif d'élargir le potentiel des données sur l'éducation grâce à la mise en relation et au traitement approfondi de sources disponibles actuellement ou qui seront collectées à l'avenir.

Describeurs: Base de données, Données de l'éducation, Entrepôt de données, Protection des données, Recherche en éducation, Système éducatif, Suisse, École

Einleitung

Bildung ist wichtig. Über diesen Punkt besteht weitgehende Einigkeit. Wie aber gelangt man zu einer guten Bildung? Und wie gelangt eine Gesellschaft zu gut gebildeten und gut ausgebildeten Bürgerinnen und Bürgern? Was gehört eigentlich dazu, und auf welchem Niveau liegt die Bildung unserer Gesellschaft? Es gibt in der Schweiz freilich diverse Studien dazu am bekanntesten ist vielleicht die von der OECD lancierte PISA-Studie, die international vergleichend alle drei Jahre die Kompetenzen von 15-jährigen Schülerinnen und Schülern in Kernbereichen der schulischen Bildung erfasst (z. B. Konsortium PISA 2019). Seit 2016 wird auch ein nationaler Test durchgeführt, die Überprüfung des Erreichens der Grundkompetenzen (ÜGK), die auf der Basis schweizerisch festgelegter Bildungsziele im kantonalen Vergleich auch Aussagen zur Harmonisierung der obligatorischen Schule erlaubt (Konsortium ÜGK 2019 a, 2019b). Auch existieren Studien über andere Altersgruppierungen, wie beispielsweise die PI-AAC-Studie (Programme for the International Assessment of Adult Competencies) der OECD.

Dies sind nur die prominentesten Beispiele von Studien, die Bildungsdaten produzieren, um mit der daraus gewonnenen Information Fragen wie die oben gestellten zu

beantworten. Gleichwohl gelingt dies nur in sehr bescheidenem Maße, und das ist verständlich, weil die Materie selbst so komplex ist. Auch wenn die genannten Studien ohne Zweifel bis zum heutigen Tag eine beachtliche Wirkung erzielen konnten, so setzen sie dennoch nur zeitweise eine Lupe an, um einen mikroskopischen Bereich des Bildungsuniversums genauer zu betrachten. Wesentlich mehr ist derzeit noch nicht möglich, auch wenn das Potenzial auf der Hand liegt und bessere Daten sowohl für die Forschung als auch für die Steuerung des Bildungssystems dringend benötigt werden.

Die aktuelle Notsituation in der Coronavirus-Krise macht das Problem ungenügender Daten besonders deutlich. Von einem Tag auf den anderen waren Lernende, Lehrpersonen und Schulverantwortliche sowie Eltern und andere Betreuungspersonen dazu gezwungen, sich neu zu organisieren und auf den täglichen Präsenzunterricht in der Schule zu verzichten. Wie das funktioniert hat und welche Auswirkungen es möglicherweise auf die Chancengerechtigkeit oder andere Aspekte des Schullebens hat und haben wird, darüber können wir mangels Daten zurzeit nicht viel mehr als spekulieren. Wir wissen nicht einmal zuverlässig, wie viele Schülerinnen und Schüler über einen eigenen Internetanschluss oder die entsprechende Hardware verfügen, geschweige denn, ob und wie häufig sie digitale Dienste nutzen und welche das sind. Wollen wir mehr über die Funktionsweise und die Qualität unseres Bildungssystems wissen und erforschen, dann brauchen wir Daten dafür.

Aber das Problem geht noch weiter. Wenn Bund und Kantone durch die Schweizer Verfassung einerseits dazu verpflichtet sind, die Qualität und die Durchlässigkeit des Bildungsraums Schweiz zu gewährleisten, und ihnen andererseits schon nur zur Beurteilung der Lage die notwendige Information fehlt, dann liegt das Problem nicht mehr nur bei der Forschung, sondern stellt zusätzlich ganz offensichtlich ein Hindernis für die Erfüllung dieses demokratischen Auftrags dar. Normalerweise bedienen sich Bund und Kantone als Grundlage für den Verfassungsauftrag an einer breiten Palette von Daten, die sie im Rahmen des Bildungsmonitorings sammeln und in dem vierjährlich erscheinenden Bildungsbericht zusammenfassen (z. B. SKBF 2018). Die Erkenntnisse aus zahlreichen wissenschaftlichen Studien ebenso wie unterschiedliche amtliche Statistiken fließen in die Fülle der Dokumente ein, aus der der Bildungsbericht schöpfen kann. Die Programmperiode des Tätigkeitsprogramms der Schweizerischen Konferenz der kantonalen Erziehungsdirektoren (EDK) und die Botschaft zur Förderung von Bildung, Forschung und Innovation (BFI-Botschaft) des Bundes sind auf diesen Zeitraum abgestimmt, damit aufgrund der Er-

kenntnisse aus dem Bildungsbericht die geeigneten Maßnahmen in den Programmen verankert werden können.

Aber von normalen Zeiten können wir im Jahr 2020 wirklich nicht sprechen. Es sind besondere Zeiten, in denen wir einen schnelleren Zugang zu Daten brauchen. Die Veränderungen kamen zu plötzlich, als dass man darauf adäquat mit einem normalen Vierjahresqualitätszyklus reagieren sollte. Was können wir also in dieser spezifischen Lage tun? Die Erhebung von wissenschaftlichen Daten als Informationsgrundlage benötigt im Normalfall einen langen Zeitraum. Die Arbeiten zu einer PISA-Studie werden beispielsweise jeweils mindestens drei Jahre vor der Erhebung und vier Jahre vor der Erstveröffentlichung der Ergebnisse begonnen. Aber können wir es uns in der jetzigen Zeit überhaupt leisten, solange zuzuwarten? Und wie könnten wir bei schnellerem wissenschaftlichen Handeln die notwendige Qualität sicherstellen? Dies sind wichtige Fragen vor allem vor dem Hintergrund, dass die klassische Datenerhebung in den Schulen bei Studien wie der ÜGK in diesem Jahr nicht möglich war. Auch der Feldtest zu PISA 2021 konnte nicht durchgeführt werden, und es steht zu befürchten, dass wir deshalb für einen etwas längeren Zeitraum keine neuen Daten aus solchen Studien zu erwarten haben.

Wenn wir diese neue Ausgangslage genau betrachten, wird relativ schnell klar, dass wir, wenn wir Wissen und Forschung aus Bildungsdaten wollen, dieses aus bereits vorhandenen Daten oder aus zukünftig anfallenden Daten erschließen müssen, und dass wir damit vom klassischen Erhebungsmodus der bisher bewährten Studien abweichen und neue Konzepte anwenden müssen. Die begonnene und nun durch die Krise weiter forcierte digitale Transformation ist in diesem Zusammenhang ein Glücksfall. Es fallen nämlich vermehrt Daten aus der Nutzung digitaler Dienste an, auch im Bildungsbereich. Die Herausforderung wird es nun sein, die verfügbaren Daten für die Forschung und letztlich auch zur Erkenntnisgewinnung für die Steuerung des Bildungssystems nutzbar zu machen. Und da sind sowohl seitens der Forschung als auch seitens der Bildungsbehörden noch einige Hürden zu nehmen. Vor allem sind Wege zu finden, wie ein erweitertes Forschungspotenzial mit einem verlässlichen Datenschutz vereint werden kann.

Virtual Educational Observatory (VEO): Nutzung vorhandener und zukünftig anfallender Bildungsdaten

Vorhandene oder zukünftig anfallende Daten für die Bildungsforschung und zur Erweiterung des Wissens über das Bildungssystem zu nutzen, ist die Grundidee des Projekts „Virtual Educational Observatory“ (VEO), das an der Fachhochschule Graubünden (FHGR) bereits vor der Coronavirus-Krise konzipiert wurde und mit einer Finanzierung des Schweizerischen Nationalfonds im Rahmen des Nationalen Forschungsprogramms (NFP77) im Mai 2020 starten konnte. Ziel des Projekts ist es, verfügbare Datensätze so zu verknüpfen, dass sie als zusammengesetztes Bild zur Wissensgenerierung im Rahmen der Bildungsforschung und des Bildungsmonitorings nutzbar werden. Wir stellen es uns vor wie bei einem Puzzlespiel: sehen wir nur einzelne Teile, ist es uns in der Regel unmöglich zu erkennen, was das Gesamtbild darstellt. Je mehr Teile wir jedoch zusammenfügen, desto klarer wird das Bild und desto genauer können wir Details und Zusammenhänge erkennen. Zudem entdecken wir auch, in welchen Bereichen noch wichtige Lücken bestehen bleiben, die uns am Erfassen des Gesamtbildes hindern.

Neben amtlichen Statistiken und Erkenntnissen aus den behördlich beauftragten Untersuchungen, gehören selbstverständlich auch alle anderen, vom wissenschaftlichen Forschungsbedarf geleiteten Erkenntnisse zu den Puzzleteilen, die wir versuchen müssen zusammenzusetzen. Der Weg zu einer reicheren Datenlandschaft führt also über das Zusammenspielen einer breiten Palette von administrativen Daten und Forschungsdaten, die beispielsweise aus Tests und Befragungen stammen. Es ist zusätzlich dazu zu prüfen, welche weiteren digitalen Datenquellen, die möglicherweise nicht explizit für die Forschung produziert wurden, eingebunden werden können (auf das Potenzial von Bildungsdaten aus der Nutzung digitaler Dienste gehen wir im nächsten Abschnitt expliziter ein).

Voraussetzung für jede Verwendung von Daten im Rahmen von VEO ist aber immer, dass alle Datenquellen über verlinkbare Inhalte verfügen, was wiederum voraussetzt, dass auch die Forschenden diese Möglichkeit von Beginn an, also bei der Konzeption ihrer Projekte, mitberücksichtigen. Im Projekt VEO wird deshalb angestrebt, die Möglichkeiten der Verlinkung von Forschungsdaten zu propagieren und zu unterstützen.

Das Schweizerische Institut für Informationswissenschaft (SII) im Themenschwerpunkt Angewandte Zu-

kunftstechnologien der FHGR, in dem das Projekt angesiedelt ist, ist darauf spezialisiert, Lösungen zu aktuellen und zukünftigen Herausforderungen im Bereich der Produktion, Organisation und Distribution von Information und Wissen zu entwickeln und zu erarbeiten. Insofern kann das Projekt VEO in idealer Weise von den vielfältigen Kompetenzen und Erfahrungen der beteiligten Wissenschaftlerinnen und Wissenschaftler des Instituts sowie ihres erweiterten Netzwerkes profitieren.

Das Projekt ist interdisziplinär angelegt, es arbeiten Fachleute aus der Bildungsinformatik, der Soziologie und dem Data Mining zusammen, um den für diese komplexe Aufgabe notwendigen mehrperspektivischen Blick zu gewährleisten. Zunächst wird dabei eine Übersicht vorhandener und zu erwartender Datensätze erstellt und jeweils die Qualität dieser Daten im Hinblick auf deren Wiederverwendung und Verknüpfbarkeit überprüft. Mittels Big-Data-Technologien und modernen maschinellen Lernverfahren werden sie dann weiterverarbeitet. Um das Potenzial und die Machbarkeit von Analysen an zusammengesetzten Daten aufzuzeigen, sollen schließlich konkrete Anwendungsfälle exemplarisch berechnet und beschrieben werden. Einer dieser Anwendungsfälle wird sich mit den Einflussfaktoren digitaler Kompetenz auseinandersetzen.

Dass der Bedarf an zusätzlichen Erkenntnissen, insbesondere über die Fragen der Chancengerechtigkeit, sich durch die Coronavirus-Krise erheblich verschärfen würde, wusste an der FHGR bei der Konzeption des Projektes natürlich noch niemand. Dass das Projekt nun aber direkt im Mai 2020 starten konnte, ist eine Chance, die jetzt genutzt werden sollte.

Das Potenzial von Bildungsdaten aus der Nutzung digitaler Dienste

Bezogen auf die Coronavirus-Krise wird es rückwirkend schwierig sein, die Situation während der Schließung der Schulen im Frühjahr 2020 empirisch zu beurteilen. Man wird sich dabei voraussichtlich im Wesentlichen auf Befragungen stützen können, wie sie bereits heute, beispielsweise aus dem Schulbarometer, zur Verfügung stehen (Huber et al. 2020). Da es aber letztlich um die Leistungsfähigkeit eines Systems und seiner Organisationen geht, ist hierbei mit einem hohen Einfluss der sozialen Erwünschtheit bei der Beantwortung der zum Teil post-hoc durchgeführten Befragungen zu rechnen. Auch die Repräsentativität der Stichproben wird nicht immer gegeben sein. Wesentlich zuverlässiger wäre es, wenn man die tat-

sächlich bei der Anwendung der digitalen Dienste angefallenen Daten auswerten könnte. Diese aber stehen ebenfalls zurzeit nicht zur Verfügung, es gibt wie oben erwähnt nicht einmal eine Angabe über die Nutzungshäufigkeit der digitalen Dienste. Aus datenschutzrechtlicher Sicht könnte man meinen, dies sei begrüßenswert, aber dieser Blick greift möglicherweise zu kurz, denn es gibt ja die Daten dennoch, und sie werden auch verwendet. Nur werden sie in der Regel ausschließlich von rein kommerziellen Diensteanbietern zur Qualitätsentwicklung ihrer Produkte, nicht aber für die Forschung und die Qualitätsentwicklung des Bildungssystems verwendet.

Diese Schräglage wurde von den Kantonen bereits vor einiger Zeit und explizit nicht coronavirusbedingt erkannt. Sie haben deshalb eine Föderation ins Leben gerufen, die über eine persönlich verwendete digitale Identität den Zugang und den Datenaustausch zwischen den in der Schule agierenden Diensteanbietern und den Nutzerinnen und Nutzern koordiniert. Auf diese Weise kann beispielsweise festgelegt werden, welche Daten an den Diensteanbieter weitergeleitet werden dürfen und welche nicht (vgl. EDK 2019). Aus datenschutzrechtlicher Sicht ist dies ein großer Fortschritt. Unter dem Namen Edulog können seit Kurzem Kantone bzw. Gemeinden auf dieses Angebot zugreifen. Es sind zurzeit erst einige Kantone, die in dieser ersten Phase teilnehmen, aber es ist zu hoffen, dass Edulog bald flächendeckend in der gesamten Schweiz verwendet wird, denn eine Föderation lebt schließlich von der Vielfalt und der großen Zahl.

Die Föderation sammelt keine Daten und verfügt deshalb auch nicht über Daten, die sie weitergeben könnte. Edulog soll ja gerade einen Schutz bieten und die Datennutzung in der Schule regulieren. Dies ist sicherlich am Besten möglich, wenn die Föderation nicht selbst über Daten verfügt. Die Überlegenheit dieser Architektur liegt auf der Hand, auch wenn wir aktuell in eher kurzsichtiger Weise geneigt sein könnten, eine direkte Datennutzung aus der Föderation heraus zu wünschen.

Warum Edulog dennoch einen wichtigen Beitrag zur Gewinnung von Erkenntnissen leisten kann in Zeiten, in denen die klassischen Datenerhebungsmethoden zumindest vorübergehend obsolet geworden sind, ist: Edulog reguliert die Datennutzung und ist dadurch in der Lage, sie transparent zu machen. Die Eignerinnen und Eigner der Daten, also Personen, die über Edulog digitale Dienste nutzen, könnten beispielsweise geregelt ihr Einverständnis dazu geben, dass ihre Daten zu Forschungszwecken stichprobenbasiert und pseudonymisiert weiterverarbeitet werden. Sie könnten selbst mitbestimmen, welche Daten das sein sollen und damit die Rechte an ihren Daten bei sich behalten. Es ist zu hoffen, dass auch von den verant-

wortlichen Stellen dieses Potenzial erkannt wird und dass bald die notwendigen Schritte unternommen werden, um mit Edulog einen Beitrag für die Forschung und letztlich für die Qualitätsentwicklung des gesamten Bildungssystems zu leisten.

Fazit

Bildungsdaten werden heute aus vielfältigen Quellen geschöpft, sei es aus Bildungsstatistiken, aus der Bildungsforschung oder digitalen Anwendungen, und aus jeder dieser Quellen können wichtige Erkenntnisse über Bildung und Lernen gezogen werden. Diese Daten bieten jedoch noch erweitertes Potenzial, wenn es uns gelingt, sie sinnvoll zu verknüpfen und sie gemeinsam sprechen zu lassen. Genau dies ist das Ziel des Projekts VEO: offiziell erhobene Datensätze (wie Bildungsstatistiken), bestehende nationale und internationale Forschungsprojekte (PISA, ÜGK, TREE etc.) und Umfragedaten mit neuen Datensätzen (insbesondere Anwendungsdaten) zusammenzubringen. Damit können wir, ohne dazu zwingend neue Daten erheben zu müssen, Fragen angehen, die bisher teilweise noch nicht zufriedenstellend beantwortet werden konnten. Dies ist heute umso bedeutender, weil die klassischen Datenerhebungen in der Schule im Frühjahr 2020 und möglicherweise bis auf Weiteres nicht stattfinden können. Wie und ob die Nutzung und Verknüpfung zukünftig anfallender Daten aus digitalen Diensten realisiert werden kann, ist jedoch noch offen. Aktuell sind wir noch weit davon entfernt, solche Überlegungen in der Praxis umsetzen zu können, aber evtl. kann VEO dazu beitragen, einen Schritt in eine Richtung zu tun, die uns ermöglicht, das Systemwissen zu vergrößern, Aufwand und Kosten von Datenerhebungen zu reduzieren und den Datenschutz besser zu gewährleisten. Damit dieses Projekt wirklich erfolgreich sein kann, benötigt es aber auch eine Unterstützung der politischen Behörden zur Öffnung des Zugangs zu den relevanten Daten.

Literatur

- EDK (2019). Organisationsreglement der Föderation der Identitätsdienste im Bildungsraum Schweiz. Bern, <https://www.edk.ch/dyn/24092.php> [20.7.2020].
- Huber, S., Gerhard, G., Paula S., Schneider, N., Helm, C., Schwander, M., Schneider, J., Pruitt, J. (2020). COVID-19 und aktuelle Herausforderungen in Schule und Bildung. Erste Befunde des Schul-Barometers in Deutschland, Österreich und der Schweiz. Münster: Waxmann.

Konsortium PISA.ch, Hrsg. (2019). PISA 2018. Schülerinnen und Schüler der Schweiz im internationalen Vergleich. Bern und Genf: SBF/EDK und Konsortium PISA.ch.

Konsortium ÜGK, Hrsg. (2019a). Überprüfung der Grundkompetenzen. Nationaler Bericht der ÜGK 2016: Mathematik 11. Schuljahr. Bern und Genf: EDK und SRED.

Konsortium ÜGK, Hrsg. (2019b). Überprüfung der Grundkompetenzen. Nationaler Bericht der ÜGK 2017: Sprachen 8. Schuljahr. Bern und Genf: EDK und SRED. ISBN 978-2-940238-28-6

SKBF, Hrsg. (2018). Bildungsbericht Schweiz 2018. Aarau: Schweizerische Koordinationsstelle für Bildungsforschung.



Dr. Vera Husfeldt
 Fachhochschule Graubünden
 Schweizerisches Institut für
 Informationswissenschaft
 Pulvermühlestrasse 57
 7000 Chur
 Schweiz
vera.husfeldt@fhgr.ch

Dr. Vera Husfeldt ist Dozentin für Digital Education / Bildungsinformatik im Institut für Informationswissenschaft der Fachhochschule Graubünden. Nach ihrem Doktorat im Bereich der Bildungsforschung an der Humboldt-Universität zu Berlin war sie zunächst als Juniorprofessorin an der Universität Göttingen und dann als Professorin für empirische Schulforschung und Leistungsmessung an der Pädagogischen Hochschule der Fachhochschule Nordwestschweiz beschäftigt. Von 2012 bis 2019 leitete sie die Abteilung Qualitätsentwicklung bei der Schweizerischen Konferenz der kantonalen Erziehungsdirektoren.



David Schiller, M. A.
 Fachhochschule Graubünden
 Schweizerisches Institut für
 Informationswissenschaft
 Pulvermühlestrasse 57
 7000 Chur
 Schweiz
david.schiller@fhgr.ch

David Schiller ist Dozent im Institut für Informationswissenschaft der Fachhochschule Graubünden. Nach einer kaufmännischen Ausbildung studierte er Soziologie und Geschichte an der Universität Erlangen-Nürnberg. Danach arbeitet er in verschiedenen Institutionen zu den Themen Bildungsverläufe, Datenmanagement, Datenschutz und internationale Forschungsnetzwerke. Sein aktuelles Forschungsinteresse liegt im Bereich InformationLearning.



Sharon Alt, MSc. BA.
 Fachhochschule Graubünden
 Schweizerisches Institut für
 Informationswissenschaft
 Pulvermühlestrasse 57
 7000 Chur
 Schweiz
sharon.alt@fhgr.ch

Sharon Alt absolvierte den Masterstudiengang Business Administration Major Information and Data Management der Fachhochschule Graubünden. Wissenschaftliche Arbeitsschwerpunkte sind Informationsmanagement, Digitale Bildung und Digitale Transformation der Bauwirtschaft.