

Big Data in der Bildung – Learning Analytics, Educational Data Mining und Co.

Tim Jülicher, Institut für Informations-, Telekommunikations- und Medienrecht (ITM),
Westfälische Wilhelms-Universität Münster

1 Digitalisierung in Bildungseinrichtungen

Die Schule, wie wir sie heute kennen, hat sich in den Jahren seit ihrer „Erfindung“ im 17. Jahrhundert (Kerstan 2013) kaum verändert. Klassenarbeiten auf Papier, Skizzen an der Tafel, die Vor- und Nachbereitung des Unterrichts mithilfe von Lehrbüchern – allen methodischen und pädagogischen Reformen zum Trotz deutete lange nichts darauf hin, dass sich an diesen Rahmenbedingungen jemals etwas ändern würde. Doch in den vergangenen Jahren hat die digitale Revolution auch das Klassenzimmer und die Hörsäle ergriffen: Schon heute finden sich in vielen Schulen Smart Boards, Vorlesungen werden durch Online-Angebote¹ begleitet und Zuhörer können sich durch Abstimmungs- und Feedback-Apps² in Echtzeit einbringen. Daneben ist der Einsatz von Tablet-Computern längst keine Seltenheit mehr.

Fest steht: Dies ist erst der Anfang, die nächste Stufe der Digitalisierung wird das Lernen und Lehren der Zukunft nachhaltig verändern. Viele technologische Neuerungen lassen sich heute noch gar nicht absehen – die Gegenwart gibt aber bereits einen Vorgeschmack auf das, was uns in nicht allzu ferner Zukunft erwarten wird. Es ist damit zu rechnen, dass die Implementierung neuer Technologien sowohl im schulischen als auch im universitären Kontext erheblich an Fahrt aufnehmen wird. Big Data und „Education 2.0“ haben das Potential, den gesamten Bildungsprozess grundlegend zu transformieren.

¹ Als namhafte und in über 220 Ländern verbreitete Open Source-Lernplattform sei hier als Beispiel Moodle (www.moodle.org) genannt.

² Beispiele sind das deutsche Open Source-Projekt „ARSnova“, die Hörsaal-App „SMILE – Smartphones in der Lehre“ oder kommerzielle Angebote wie „Top Hat“.

Auf einen Blick: Big Data in der Bildung

- Ausgangspunkt für Big Data-Anwendungen im Bildungsbereich ist die zunehmende Integration technischer Geräte und digitaler Ressourcen in den schulischen und universitären Alltag. Dazu zählen vor allem spezifische Online-Angebote (z.B. Blended Learning, eLearning und Social Media).
- Die Vielfalt der gesammelten Daten gewährt umfassende Einblicke in das Lernverhalten von Klassen und einzelnen Schülern. Die Technologien dahinter sind als Educational Data Mining (EDM) und Learning Analytics (LA) bekannt und versprechen individuelles Lern-Management, adaptives Lernen und vieles mehr.
- In Europa ist Big Data & Education bislang noch kein großes Thema – anders dagegen in den USA. Hier wird bereits über das Potential und die Risiken der breit angelegten Verknüpfung und Auswertung von Bildungsdaten diskutiert.
- Herausforderungen: Der Einsatz von Bildungs-Apps, Classroom-Management-Systemen und Online-Angeboten ist bislang weitgehend unreguliert. Probleme drohen nicht nur in datenschutzrechtlicher Hinsicht. Die Analyse personenbezogener Bildungsdaten wirft auch ethische Fragen auf (z.B. bei einer bloßen Einordnung eines Schülers anhand seines prognostizierten Lernverhaltens). Schließlich sind Bildungsdaten von besonderem wirtschaftlichem Wert.

2 Die Zukunft der Bildung – Szenarien, Akteure und Potentiale

Lange Zeit beschränkte sich die Analyse des individuellen Lernverhaltens auf mehr oder minder differenzierte Leistungsbewertungen wie Klausurnoten, Zeugnisse oder Scheine zum Ende eines Schuljahres oder Semesters.

Schon heute ermöglicht der Einsatz neuer Technologien aber ein viel detaillierteres Bild des Einzelnen und der Masse: Ein Dozent kann auf E-Learning-Plattformen nicht nur einsehen, wie oft die Folien der letzten Veranstaltung heruntergeladen wurden, sondern erhält zugleich Statistiken darüber, wie lange sich Studenten einloggen und womit sie sich währenddessen beschäftigten.³

Bei Massive Open Online Courses (MOOCs) kann der „Clickstream“ der Teilnehmer verfolgt und so präzise auf das jeweilige Lernverhalten und etwaige Lernprobleme geschlossen werden (Breslow et al., 2013; Picciano, 2012). Dort, wo Tablets, Smartphones und E-Books klassische Schulbücher ersetzen, lässt sich ohne weiteres nachvollziehen, wie schnell Schüler lesen, welche Stellen sie zur Klausurvorbereitung wiederholen oder wer die Pflichtlektüre gar nicht gelesen hat.

Diese Beispiele machen deutlich, dass durch die Implementierung neuer Technologien und Konzepte wie des integrierten Lernens (sog. Blended Learning)⁴ Unmengen an Daten generiert werden – sei es zum Lernverhalten einzelner Schüler, dem Lernfortschritt der gesamten Klasse oder der Akzeptanz des Unterrichtskonzepts. Tatsächlich liegen die Herausforderungen und Verheißungen von „Education 2.0“ aber erst im nächsten Schritt, nämlich der Verknüpfung und Analyse dieser vielfältig aggregierten Daten.

3 Educational Data Mining und Learning Analytics

Hier kommen **Educational Data Mining** (EDM) und **Learning Analytics** (LA) ins Spiel. Eine erste Herausforderung stellt es dar, die im Zuge des „technologisierten“ Lernens anfallenden Informationen zu organisieren. Neben strukturierten Daten wie der IP-Adresse oder dem Benutzernamen des Lernenden sind dies zunehmend auch unstrukturierte Daten wie Texte aus Diskussionsforen oder Video- und Audiodateien. Hinzu kommen Metadaten wie Aktivitätszeiten oder inhaltliche Zusammenhänge.

EDM ermöglicht es, diese Informationen zu ordnen und in Zusammenhang zu bringen, um sie für weitere Analysen und Prognosen fruchtbar zu machen (ausführlich Ebner & Schön, 2013). Insoweit ähnelt das EDM dem „kommerziellen“ Data Mining (d.h. der systematischen Auswertung großer Datenbestände mit dem Ziel, neue, vor allem ökonomisch wertvolle, Erkenntnisse zu gewinnen), wie man es aus der Werbe- und Finanzwirtschaft kennt.

Derweil hat Learning Analytics die Interpretation der gesammelten Daten zum Ziel (zum Verhältnis von EDM und LA zueinander Siemens & Baker, 2012). Im Kern geht es darum, den individuellen Lernprozess zu optimieren: Dies geschieht nicht nur durch eine umfassende Abbildung des bisherigen Lernverhaltens, vielmehr sollen aus dem vorhandenen Datenbestand zugleich Erkenntnisse und (wahrscheinlichkeitsbasierende) Prognosen für die Zukunft abgeleitet werden, sog. **predictive analytics** (Siemens, 2010).

So können beispielsweise unerkannte Lernschwächen eines Schülers frühzeitig und gezielt adressiert werden, um ihn auf die nächste Klassenarbeit vorzubereiten.

4 Akteure

Wer profitiert von den versprochenen Verheißungen? Hier ist zunächst an Lernende und Lehrende zu denken. Die Lehrkraft kann sowohl das Lernverhalten der ganzen Klasse als auch den Fortschritt einzelner

³ Dazu ausführliche Informationen bei <https://docs.moodle.org/28/de/Kursberichte>.

⁴ Gemeint ist die didaktische Verknüpfung klassischer Präsenzveranstaltungen und moderner E-Learning-Inhalte zu einem einheitlichen Lehr- und Lernkonzept.

Schüler verfolgen und individuell darauf eingehen. Zugleich erhält der Dozent ein unmittelbares persönliches Feedback. Schüler werden ihrem Leistungsniveau und -potential entsprechend gefordert und gefördert.

Daneben können Wissenschaft und Verwaltung von einer Auswertung der im Bildungskontext gesammelten Daten profitieren, beispielsweise durch die Evaluation von Einrichtungen, Dozenten, Lehrplänen oder Studienprofilen. In diesem Zusammenhang sind auch die Interessen von Herstellern und Entwicklern entsprechender Analysetools und -systeme zu berücksichtigen.

Schließlich ist zu erwarten, dass auch die Wirtschaft ein großes Interesse an den gesammelten Daten hat. Dies gilt nicht nur für klassische Dienstleister im Bildungsbereich (z.B. Fachverlage oder Nachhilfeinstitute), sondern auch Arbeitsvermittler (Jobcenter, Agentur für Arbeit) und Recruiting-Abteilungen.

5 Datenquellen – Der Schüler als Datenlieferant?

Während sich in anderen Lebensbereichen bereits klar abzeichnet, wie und wo Big Data zum Einsatz kommen kann (z.B. in der Wirtschaft – Stichwort **Industrie 4.0**), sind die Anwendungsbereiche im Bildungskontext bislang weniger klar umrissen. Fest steht jedoch, dass Schüler, Lehrer, Studenten und Dozenten schon heute eine Unmenge an Daten generieren, deren Analyse großes Potential und (unabsehbare) Risiken birgt. Doch um welche Daten geht es eigentlich?

Anhaltspunkte liefert ein Blick in die USA, wo die Digitalisierung an Schulen und Universitäten deutlich weiter vorangeschritten ist: Hier greifen Lehrer zunehmend zu Classroom-Management-Systemen oder Apps, die den Unterricht bereichern können. Tatsächlich sind die wenigsten dieser Anwendungen jedoch von staatlichen Aufsichtsbehörden genehmigt. Ihr Einsatz erfolgt weitgehend unreguliert und es fehlt an IT-Sicherheitsstandards. Auch die Qualität der Datenerhebung ist höchst unterschiedlich.

Weitere Datenquellen bilden eLearning-Angebote, von Schulen zur Verfügung gestellte Laptops oder Tab-

lets sowie Prepaid-Mensakarten oder mit RFID-Chips versehene Schülerschulenausweise (Hill, 2014).

Die so erhobenen Daten landen nicht nur bei Schulen, sondern auch bei Unternehmen. So erklärt ein Anbieter adaptiver Lernsysteme, ihm lägen weitaus mehr Daten über Schüler vor als jeder Schule oder Google (Ferreira, 2012).

Primärer Datenlieferant ist bei alledem der Schüler selbst. Er gibt in großem Umfang – und bisweilen unfreiwillig – personenbezogene und hochsensible Daten preis, was die Angst vor dem „gläsernen Schüler“ schürt und Datenschützer auf den Plan ruft.

6 Herausforderungen

Big Data verspricht sowohl im universitären als auch im schulischen Bereich revolutionäre Veränderungen. Die Entwicklung in den USA hat gezeigt, dass vor allem Universitäten eine Vorreiterrolle einnehmen (Slade & Prinsloo, 2013: „Higher education cannot afford to not use data“).

Hierzulande steht die Entwicklung dagegen noch am Anfang. Dies bietet die Chance – frühzeitig und im Dialog mit allen Beteiligten – tragfähige Strategien im Umgang mit der Verknüpfung, Auswertung und Analyse von Bildungsdaten zu entwickeln.

Herausforderungen ergeben sich insbesondere mit Blick auf eine transparente und datenschutzkonforme Umsetzung, die den Risiken der fortschreitenden Datenerhebung im Bildungskontext gerecht wird. Hier muss ein sicherer und effektiver Schutz der Daten aller Betroffenen gewährleistet sein.

Schließlich gilt es zu bedenken, dass die Schaffung einer den gesamten Bildungsweg durchziehenden, personenbezogenen Datensammlung erhebliche wirtschaftliche, pädagogische und ethische Konsequenzen mit sich bringt. Die damit einhergehenden Risiken sind zum jetzigen Zeitpunkt kaum abschätzbar.

Literaturnachweise

- Kerstan, T. (2013). Wer hat die Schule erfunden? *Die ZEIT* 25/2013.
- Breslow, L., Pritchard, D. E., DeBoer, J., Stump, G. S., Ho, Seaton, D. T. (2012). Studying Learning in the Worldwide Classroom. *Research into edX's First MOOC. Research & Practice in Assessment* 8, 13-25.
- Picciano, A. G. (2012). The Evolution of Big Data and Learning Analytics in American Higher Education. *Journal of Asynchronous Learning Networks* 16 (3), 9-20.
- Ebner, M., & Schön, M. (2013). Das Gesammelte interpretieren – Educational Data Mining und Learning Analytics. In M. Ebner & S. Schön (Hrsg.), *Lehrbuch für Lernen und Lehren mit Technologien*. Epubli. Online verfügbar unter <http://l3t.tugraz.at/index.php/LehrbuchEbner10/article/download/119/117>.
- Siemens, G., & Baker, R. S. (2012). Learning Analytics and Educational Data Mining: Towards Communication and Collaboration. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*. 252-254.
- Siemens, G. (2010). *What Are Learning Analytics?*. Online verfügbar unter <http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/>.
- Hill, A. (2014). A day in the life of a data mined kid, Marketplace September 15. Online verfügbar unter <http://www.marketplace.org/topics/education/learningcurve/day-life-data-mined-kid>.
- Ferreira, J. (2012). Knewton – Education Datapalooza. Online verfügbar unter <https://www.youtube.com/watch?v=Lr7Z7ysDluQ>.
- Slade, S., Prinsloo, P. (2013). Learning Analytics: Ethical Issues and Dilemmas. *American Behavioral Scientist* XX (X), 1-20. doi: 0.1177/0002764213479366

Vertiefungshinweise: Literatur und Links

- Daniel, K., & Butson, B. (2016). *Big Data and Learning Analytics in Higher Education*. Berlin: Springer. (im Erscheinen).
- Ebner, M., & Schön, S. (Hrsg.). (2013). *Lehrbuch für Lernen und Lehren mit Technologien*. Epubli. Online verfügbar unter www.l3t.eu.
- Hill, A. (2014). A day in the life of a data mined kid, Marketplace, September 15. Online verfügbar unter <http://www.marketplace.org/topics/education/learningcurve/day-life-data-mined-kid>.
- Mayer-Schönberger, V., & Cukier, K. (2014): *Lernen mit Big Data. Die Zukunft der Bildung*. München: Redline Verlag.
- Reyes, J. A. (2015). The Skinny on Big Data in Education: *Learning Analytics Simplified. TechTrends* 59 (2), 75-80.
- UCL Big Data Institute. (2014). Konferenz. Big Data and Education: What's the Big Idea? www.ucl.ac.uk/public-policy/events/big-data-and-education.



ABIDA (Assessing Big Data) **Über die Dossiers**

Das Projekt ABIDA, gefördert vom Bundesministerium für Bildung und Forschung, lotet gesellschaftliche Chancen und Risiken der Erzeugung, Verknüpfung und Auswertung großer Datenmengen aus und entwirft Handlungsoptionen für Politik, Forschung und Entwicklung. Dabei nähert ABIDA sich dem Thema Big Data aus einer grundlegend interdisziplinären Perspektive. Mehr Informationen finden Sie auf www.abida.de.

In den ABIDA-Dossiers werden regelmäßig ausgewählte Big Data-Themen kurz und prägnant dargestellt, um dem Leser einen Überblick zu liefern und einen Einstieg in die Thematik zu ermöglichen. Weitere Dossiers sind verfügbar unter www.abida.de/content/dossiers.