

# **Big Data – Big Challenges – Big Chances: Datenmanagement in den Informatikunterricht!**

Andreas Grillenberger

Friedrich-Alexander-Universität Erlangen-Nürnberg  
Didaktik der Informatik, Martensstr. 5a, 91058 Erlangen  
andreas.grillenberger@fau.de

**Abstract:** Die strukturierte Speicherung von Informationen stellt ein zentrales Thema in der Informatik und damit auch im Informatikunterricht dar. Durch die zunehmende Bedeutung von Big Data findet in diesem Bereich zurzeit ein Paradigmenwechsel statt, der sich gleichzeitig auch auf den Alltag auswirkt. Sowohl die fachlichen als auch die gesellschaftlichen Aspekte von Big Data haben dabei starken Einfluss auf den Unterricht. Insbesondere müssen die Inhalte im allgemeinbildenden Schulunterricht verschiedenen Kriterien genügen, durch Innovationen wie NoSQL-Datenbanken werden jedoch Konzepte, die bisher als grundlegend angesehen wurden, in Frage gestellt.

In diesem Beitrag wird die Relevanz von Big Data für die Unterrichtsgestaltung gemeinsam mit den Herausforderungen und Möglichkeiten, die sich durch die Einflüsse von Big Data ergeben, analysiert. Dadurch soll die Grundlage für eine zukunftssichere Gestaltung zukünftiger Lehrpläne zum Thema Datenmanagement geschaffen werden.

## **1 Einleitung**

„*Big Data ist ein Ausbildungsthema*“ – obwohl sich dieses Zitat der Gesellschaft für Informatik [GI 13] speziell auf die berufliche (Aus-)Bildung bezieht, hat das Thema auch deutliche Einflüsse auf den Informatikunterricht an allgemeinbildenden Schulen: Die Aufbereitung und Speicherung von Informationen stellt, neben der durch die Öffentlichkeit oft als Kernthema der Informatik betrachteten Softwareentwicklung, eine zentrale Aufgabe sowohl in der Informatik [CS06, Artikel „Informatik“] als auch im Informatikunterricht dar. Während sich bereits verschiedene Paradigmenwechsel in der Softwareentwicklung auch auf den Unterricht ausgewirkt haben – zuletzt der Wechsel zur objektorientierten Programmierung – findet ein solcher Wandel aktuell auch im Datenmanagement statt [MC13]. Dieser Paradigmenwechsel wird dabei insbesondere durch die Verarbeitung großer, unstrukturierter und schnell variierender Datenmengen – Big Data – beeinflusst.

Durch diese Entwicklungen wird sowohl die Informatik als auch der Informatikunterricht vor verschiedene Herausforderungen gestellt. Gleichzeitig ergibt sich aber auch die Möglichkeit, neue Anwendungsgebiete der Informatik darzustellen und den Informatikunterricht motivierend und modern zu gestalten. Durch eine Analyse dieser Herausforderungen und Möglichkeiten, sowie der Relevanz von Big Data für den Informatikunterricht, wird eine Grundlage für zukünftige Lehrpläne zum Thema Datenmanagement im Informa-

tikunterricht geschaffen, die einen zukunftssicheren und aktuellen Unterricht ermöglicht.

## **2 Aktueller Forschungsstand und Informatikunterricht**

Die Themen Datenbanken und Datenverwaltung/Datenmanagement sind in nahezu allen Informatiklehrplänen in Deutschland, aber auch in den GI-Empfehlungen zu Bildungsstandards in der Informatik [Puh08], obligatorisch verankert. Trotz der aktuellen Entwicklungen in den Gebieten Datenbanken und Datenmanagement herrscht in den letzten Jahren jedoch nahezu Stillstand in der fachdidaktischen Diskussion zu diesen Themen, seit zu Beginn der 1990er Jahre die Relevanz von relationalen Datenbankmanagementsystemen (RDBMS) als Unterrichtsthema ausführlich diskutiert wurde (z. B. [Wit94]).

Dies führt zu einer deutlichen Einigkeit bei der Setzung von Schwerpunkten im Datenbankunterricht über alle Bundesländer hinweg, da kaum widersprüchliche Positionen zu diesem Thema existieren. Die wichtigsten inhaltlichen Schwerpunkte im Unterricht stellen daher der Entwurf von relationalen Datenmodellen sowie SQL-Abfragen dar (z. B. [ISB09b]), zur Datenverwaltung wird üblicherweise auf RDBMS zurückgegriffen. Als Beispiele werden dazu oft kleine und fiktive Datenbanken (beispielsweise „Sportverein“ [ISB07]) betrachtet, nur in seltenen Fällen wird mit größeren Datenmengen gearbeitet. Es ist daher absehbar, dass diese Grundlage alsbald ihre Tragfähigkeit verlieren wird.

## **3 Relevanz von Big Data für den Informatikunterricht**

Um für den Unterricht relevante Inhalte festzulegen, können verschiedene Kriterienkataloge zur Auswahl allgemeinbildender Unterrichtsinhalte herangezogen werden, beispielsweise die Kriterien für fundamentale Ideen der Informatik (nach Schwill [Sch93]) oder für Great Principles of Computing (nach Denning [Den03]). Ein wichtiges Kriterium stellt dabei üblicherweise die Forderung nach einer langfristigen Relevanz der Unterrichtsinhalte dar. Dadurch soll verhindert werden, dass kurzlebiges Wissen vermittelt wird, das die Lernenden später kaum in Alltag und/oder Beruf anwenden können. Dieses Kriterium wird nach Schwill als „Zeitkriterium“, bei Denning als „recurrent“ bezeichnet. Während die bisherigen Themen im Datenbankunterricht im Sinne dieser Kriterien als für den Unterricht relevant angesehen wurden, müssen solche Bewertungen im Kontext aktueller Entwicklungen neu getroffen werden: Bisher konnte beispielsweise die Vermeidung von Redundanzen und Inkonsistenzen bei der Speicherung von Daten in normalisierten RDBMS als zentrale Idee der Datenspeicherung aufgefasst werden [EN09]. In nicht-relationalen NoSQL<sup>1</sup>-Datenbanken wird die redundante Datenspeicherung hingegen häufig eingesetzt, um das Antwortverhalten zu verbessern. Damit werden jedoch Inkonsistenzen in Kauf genommen [EFH<sup>+</sup>11]. Die Normalisierung von Datenbeständen stellt damit ein Konzept von RDBMS dar, das kaum Bedeutung in NoSQL-Datenbanken hat und dessen zukünftige Relevanz daher kaum vorhergesagt werden kann. Die Ideen von Redundanz und Inkonsistenz sind

---

<sup>1</sup> „NoSQL“ wird im Sinne von „Not Only SQL“ als Oberbegriff für nicht-relationale DBMS verwendet.

mithin für den Unterrichtskontext neu zu bewerten. Andere, oft übergeordnete, Konzepte bleiben hingegen auch in Bezug auf NoSQL-Datenbanken weiterhin relevant, z. B. Datensicherheit, Datenschutz und die strukturierte Datenspeicherung selbst, obwohl NoSQL-Datenbanken statt eines expliziten Datenschemas, wie es bei RDBMS üblich ist, oft ein implizites Schema verwenden. Anhand dieser Beispiele zeigt sich, dass die bisher im Unterricht vermittelten Inhalte in Bezug auf ihre zukünftige Relevanz neu bewertet und Schwerpunkte neu gesetzt werden müssen, um einen zeitgemäßen, dem Stand der Wissenschaft entsprechenden und zugleich zukunftssicheren Informatikunterricht zu gewährleisten.

## 4 Chancen und Möglichkeiten

Trotz der Herausforderungen, vor welche der Unterricht durch Big Data gestellt wird, werden zugleich auch vielfältige neue Möglichkeiten für die Unterrichtsgestaltung eröffnet. Indem die Einflüsse von Big Data auf den Alltag im Unterricht betrachtet werden, trägt dieser auch wesentlich zu einem besseren Verständnis gesellschaftlich relevanter Themen bei. Beispiele für solche Themen stellen dabei u. a. die Vorratsdatenspeicherung oder Überwachungsprogramme von Geheimdiensten, z. B. das PRISM-Programm der NSA, dar.

Gleichzeitig wird auch ein Bewusstsein für den Wert von Daten geschult, die auf den ersten Blick wertlos erscheinen: Während bisher im Datenschutz insbesondere auf persönliche Daten wie Namen, Adressen und Geburtsdaten geachtet wird, muss im Kontext von Big Data auch der Wert von unscheinbaren Daten beachtet werden. Solche Daten werden beispielsweise beim Besuch von Webseiten automatisch im Hintergrund übertragen. Während bisher zum Schutz persönlicher Daten oft das Prinzip der Datensparsamkeit ausreicht, ist das daher nicht mehr der Fall, wenn Big-Data-Analysen möglich sind. Es muss an dieser Stelle also die Beurteilungskompetenz der Lernenden geschult werden, so dass jeder basierend auf dem erworbenen Wissen selbst entscheiden kann, ob – in Anbetracht des Nutzens einer Anwendung – eine Datensammlung durch diese in Kauf genommen wird.

Auch durch aktuelle Entwicklungen am Arbeitsmarkt, insbesondere durch die Entstehung neuer und die Neuausrichtung bestehender Berufe, werden Anforderungen an den Informatikunterricht gestellt, da der allgemeinbildende Schulunterricht auch die Berufswahl der Schülerinnen und Schüler vorbereiten soll (vgl. z. B. [ISB09a]). Dies geschieht insbesondere, indem Einblicke in die Tätigkeiten verschiedener Berufsgruppen gewährt werden – eine Möglichkeit, die sich auch im Informatikunterricht bietet: Während die Tätigkeiten von Programmierern und auch von Datenbank-Administratoren im Unterricht meist deutlich werden, trifft dies oft nicht auf neuere informatische und informatiknahe Berufe zu. Ein Beispiel für einen solchen Beruf stellt der „Data Scientist“ [DP<sup>+</sup>12] dar: Dieser Beruf verbindet informatische, mathematische und statistische Aspekte und stellt somit auch die Vielfältigkeit und interdisziplinäre Bedeutung der Informatik dar. Indem Tätigkeiten solcher Berufsgruppen im Informatikunterricht dargestellt werden, wird außerdem gleichzeitig mit der Vorbereitung der Berufswahl auch der Blick auf das Fach Informatik geschärft: Während in der Öffentlichkeit Informatiker häufig mit Programmierern gleichgesetzt werden, wird durch die Betrachtung beispielsweise des Data Scientist deutlich, dass die Tätigkeiten eines Informatikers deutlich vielfältiger als reine Programmierung sein können.

## 5 Ausblick

Wie dargestellt bieten sich durch die Einflüsse von Big Data und NoSQL vielfältige Chancen für den Informatikunterricht – aber auch Herausforderungen, die in diesem Zusammenhang gelöst werden müssen. Für eine mögliche Neugestaltung des Unterrichts zum Thema „Datenmanagement“ müssen diese ausführlich analysiert und diskutiert werden, damit der Informatikunterricht – gerade auch durch die Betrachtung von modernen Themen – motivierend und auf einem aktuellen Stand bleibt und somit die Begeisterung von Schülerinnen und Schüler für Informatik und informatiknahe Berufe frühzeitig geweckt werden kann.

Die im Rahmen dieses Beitrags dargestellten Einflüsse, Herausforderungen und Möglichkeiten werden durch den Autor weiter untersucht werden. Unter anderem ist derzeit eine detailliertere Analyse der Auswirkungen auf die im Informatikunterricht zum Thema Datenmanagement vermittelten Konzepte in Arbeit.

## Literatur

- [CS06] V. Claus und A. Schwill. *Duden - Informatik A - Z: Fachlexikon für Studium, Ausbildung und Beruf*. Dudenverlag, Bibliographisches Institut & F.A. Brockhaus AG, 2006.
- [Den03] Peter J. Denning. Great Principles of Computing. *Commun. ACM*, 46(11):15–20, 2003.
- [DP<sup>+</sup>12] Thomas H Davenport, DJ Patil et al. Data scientist: the sexiest job of the 21st century. *Harvard business review*, 90(10):70–77, 2012.
- [EFH<sup>+</sup>11] Stefan Edlich, Achim Friedland, Jens Hampe, Benjamin Brauer und Markus Brückner. *NoSQL*. Hanser, Carl GmbH + Co., 2011.
- [EN09] Elmasri, Ramez A. und Navathe, Shamkant B. *Grundlagen von Datenbanksystemen*. Pearson Deutschland GmbH, München, 3. aktualisierte auflage. Auflage, 2009.
- [GI 13] GI (Gesellschaft für Informatik e.V.). Handlungsempfehlungen an die politischen Akteure (Big Data Days). <http://www.gi.de/fileadmin/redaktion/Hauptstadtbuero/Handlungsempfehlungen.pdf>, 2013. zuletzt überprüft: 05.03.2014.
- [ISB07] ISB (Staatsinstitut für Schulqualität und Bildungsforschung). Informatik am Naturwissenschaftlich-technologischen Gymnasium, Jahrgangsstufe 9 (Handreichung), 2007.
- [ISB09a] ISB (Staatsinstitut für Schulqualität und Bildungsforschung). Das Gymnasium in Bayern (Ebene 1 des Lehrplans des achtjährigen Gymnasiums in Bayern), 2009.
- [ISB09b] ISB (Staatsinstitut für Schulqualität und Bildungsforschung). Lehrplan für das Gymnasium in Bayern, Fach Informatik (NTG), 2009.
- [MC13] Viktor Mayer-Schönberger und Kenneth Cukier. *Big Data - Die Revolution, die unser Leben verändern wird*. FinanzBuch Verlag, München, 2013.
- [Puh08] Hermann Puhlmann et al. Grundsätze und Standards für die Informatik in der Schule: Bildungsstandards Informatik für die Sekundarstufe I. *LOG IN*, 150/151, 2008.
- [Sch93] Andreas Schwill. Fundamentale Ideen der Informatik. *Zentralblatt für Didaktik der Mathematik*, 1993.
- [Wit94] Helmut Witten. Datenbanken - (k)ein Thema im Informatikunterricht? *LOG IN*, 2, 1994.