

## **Schlüsselkompetenzen im Umgang mit Daten: Ein Thema für den Informatikunterricht**

Andreas Grillenberger und Ralf Romeike<sup>1</sup>

**Abstract:** Daten stellen in der Informatik und im Informatikunterricht ein zentrales Thema dar, das jedoch auch im täglichen Leben deutlich an Bedeutung gewinnt. *Big Data* und *Datenanalysen* werden dabei auch im gesellschaftlichen Diskurs immer stärker diskutiert, da deren Einflüsse auch abseits von Informatik und Wirtschaft heute deutlich wahrnehmbar werden. Um diese Diskussionen verstehen und daran teilhaben zu können, aber auch zum verantwortungsbewussten und selbstbestimmten Umgang mit eigenen und fremden Daten, werden vielfältige Grundkenntnisse benötigt. Daher stellt sich die Frage, welche Kompetenzen heute jeder in diesem Bereich erwerben muss. In diesem Beitrag verdeutlichen wir daher die Relevanz der Themen *Datenmanagement* und *Big Data* und erläutern auf dieser Basis eine Auswahl an Schlüsselkompetenzen aus diesem Bereich.

**Keywords:** Datenverwaltung, Datenmanagement, Schlüsselkompetenzen, Big Data, NoSQL, Datenbanken, Datenschutz, Datensicherheit, Datenanalyse, Herausforderungen

### **1 Einleitung**

Daten nehmen heute eine Schlüsselrolle im Leben ein: Riesige Datenmengen werden täglich erzeugt und verwaltet, u. a. bei der Nutzung von sozialen Medien oder von elektronischen Fahrkarten im Öffentlichen Personennahverkehr. Der Umgang mit Daten hat heute viele Facetten: Daten werden unabsichtlich oder absichtlich erzeugt, lokal oder in der Cloud gespeichert und sind unterschiedlich strukturiert. Da Daten nahezu überall erfasst werden, können durch Datenanalysen Informationen über große Teile des (Privat-)Lebens rekonstruiert werden, z. B. Gewohnheiten, Hobbies u. Ä., oft mit hoher statistischer Signifikanz. Daten haben daher einen deutlichen und weiterhin ansteigenden Einfluss auf das tägliche Leben. Ein zentraler Aspekt ist der verantwortungsbewusste Umgang mit Daten: Der Schutz eigener und fremder Daten vor Manipulation, Verlust und missbräuchlicher Nutzung gewinnt immer stärker an Bedeutung, aber auch Methoden zur Sicherstellung der Authentizität von Daten. Während diese Themen bisher hauptsächlich für Unterricht und Beruf als relevant angesehen wurden, betreffen sie nun das gesamte Leben. Informatikunterricht konzentriert sich bisher jedoch hauptsächlich auf Datenbanken, während andere Aspekte wie Datensicherheit, Datenschutz oder Metadaten außen vor bleiben oder nur am Rand betrachtet werden [GR14a]. Um den Anforderungen des täglichen Lebens zu genügen, muss sich dieser Schwerpunkt hin zur Förderung von Kompetenzen, die heute jeder in diesem Bereich benötigt, verschieben. In diesem Beitrag skizzieren wir eine Auswahl an Schlüsselkompetenzen, die verschiedene Aspekte von Datenmanagement betreffen<sup>3</sup>.

---

<sup>1</sup> Friedrich-Alexander-Universität Erlangen-Nürnberg, Didaktik der Informatik, Martensstr. 3, 91058 Erlangen  
andreas.grillenberger@fau.de, ralf.romeike@fau.de

<sup>3</sup> Eine ausführlichere Darstellung kann dem englischsprachigen Beitrag [GR14b] entnommen werden.

---

## 2 Schlüsselkompetenzen im Umgang mit Daten

**Datenspeicherung** Das Speichern von Daten ist heute eine alltägliche Aufgabe: Große Datenmengen werden, u. a. in Form von Textdokumenten, Musik oder E-Mails, gespeichert und dabei verschiedene Möglichkeiten genutzt, die sich insbesondere hinsichtlich des Speicherortes, der Art der Speicherung und der Datenstrukturierung unterscheiden. Da sich Anforderungen an die Datenspeicherung, die Nutzung von Werkzeugen wie Medienbibliotheken aber auch den Zugriff auf Daten jedoch von Fall zu Fall unterscheiden, muss man heute regelmäßig die Frage beantworten: „*Welche Daten speichere ich wo und wie ab?*“. Allein das Speichern von Daten fordert daher zumindest implizit Wissen über verschiedene Möglichkeiten der Datenspeicherung und deren Unterschiede. Um solche Entscheidungen fundiert treffen zu können, ist es daher notwendig, dass Schülerinnen und Schüler verschiedene **Möglichkeiten zur Datenspeicherung kennen und nutzen können**.

**Metadaten** Zusätzlich zu den eigentlichen Daten werden auch vielfältige weitere Informationen in Form von Metadaten gespeichert, u. a. Dateiattribute wie der Zeitstempel der letzten Änderung oder der Dateiname, aber auch Informationen wie markierte Änderungen in einem Office-Dokument. Die Menge und Ausprägung dieser Informationen hängt stark von der genutzten Anwendung und dem Dateiformat ab. Obwohl Metadaten auch sensible und nicht zur Veröffentlichung bestimmte Informationen enthalten können, werden sie oft kaum beachtet und daher beispielsweise vertrauliche Informationen unabsichtlich veröffentlicht. Beim Umgang mit Daten muss sich daher jeder **bewusst sein, dass zusätzliche Daten in Form von Metadaten in Dateien/Datensätzen enthalten sein können**. Andererseits vereinfachen Metadaten häufig auch den Umgang mit Daten: Sie tragen beispielsweise zu einer besseren Organisation und Auffindbarkeit von Dateien bei, indem sie z. B. die Suche nach Fotos anhand des Ortes, von Kategorien oder von auf dem Bild markierten Personen ermöglichen – wenn entsprechende Metadaten vorhanden sind. **Den Sinn und die Möglichkeiten von Metadaten zu verstehen und sie gezielt einsetzen zu können**, stellt damit eine weitere Schlüsselkompetenz dar.

**Redundanz und Konsistenz** Bei der Verwaltung von Daten besteht häufig die Gefahr, dass Redundanzen und Inkonsistenzen entstehen. Um die daraus resultierenden Probleme zu verhindern, werden Redundanzen meist vermieden, wie es auch im Informatikunterricht im Kontext der Datenbanksysteme üblich ist (*Normalisierung*). Um die zur Verfügung stehenden Möglichkeiten geeignet einsetzen zu können, muss daher **das Prinzip der Redundanz sowie deren Risiken für die Konsistenz von Daten verstanden werden und Möglichkeiten zur Vermeidung von Inkonsistenz gezielt eingesetzt werden**. Auch Datensynchronisation ist heute oft ein Auslöser für inkonsistente Daten, wenn Daten auf verschiedenen Geräten nicht nur gelesen, sondern auch bearbeitet werden. Nur ein Teil der dabei entstehenden Inkonsistenzen kann automatisch aufgelöst werden, in anderen Fällen muss der Nutzer manuell eingreifen. Eine weitere grundlegende Kompetenz beim Umgang mit Daten ist es daher, die **Konsequenzen von Datensynchronisation zu verstehen und mit Synchronisationskonflikten umzugehen**. Obwohl Redundanz und Inkonsistenz meist hinsichtlich ihrer Nachteile betrachtet werden, gibt es jedoch auch Anwendungsszenarien,

in denen diese beiden Konzepte gezielt eingesetzt werden. Datensicherungen stellen redundante Kopien dar und werden früher oder später inkonsistent zu den Originaldaten. Das Wiederherstellen von Daten und die Rücknahme von Änderungen wird dabei nur durch die Nutzung von Redundanz und Inkonsistenz ermöglicht. Auch beim Erstellen von Backups existieren verschiedene Prinzipien: zeitlich kontinuierlich oder diskrete Backups, inkrementell, differentiell oder komplett, online oder offline. Diese Varianten unterscheiden sich z. B. in der Ressourcennutzung und der Geschwindigkeit. Die Entscheidung, welche Variante genutzt wird, muss für jeden Anwendungsfall gezielt getroffen werden und beeinflusst die Datensicherheit deutlich. Eine weitere Schlüsselkompetenz in diesem Bereich ist daher, *redundante Daten gezielt zu Backupzwecken einsetzen zu können*.

**Datensicherheit und Verschlüsselung** Daten werden heute häufig nicht nur lokal auf dem eigenen PC gespeichert, sondern auch über ungeschützte Kommunikationskanäle wie das Internet übertragen oder auf USB-Sticks gespeichert, meist ohne geeignete Schutzmaßnahmen zu ergreifen. Es werden zwar Möglichkeiten zur Absicherung von Geräten, z. B. per Passwort oder Fingerabdruck, genutzt, wodurch auch ein Bewusstsein für die Risiken für Privatsphäre und Datensicherheit gezeigt wird. Diese Maßnahmen erhöhen die Sicherheit jedoch nur, indem der unbefugte Zugriff auf die Daten erschwert wird, während die eigentlichen Daten nicht geschützt werden: Diese können weiterhin aus dem Datenspeicher ausgelesen oder während der Übertragung mitgeschnitten werden. Ein nach dem Stand der Technik zuverlässiger Schutz kann jedoch nur durch Verschlüsselung der Datenspeicher und Kommunikationskanäle gewährleistet werden. Weitere wichtige Kompetenzen aus diesem Bereich sind daher, *den Unterschied zwischen dem Schutz eines Gerätes (oder Dienstes) und der eigentlichen Daten zu verstehen* sowie *die eigenen Daten und eigene Kommunikation zu verschlüsseln, um unbefugten Zugriff darauf zu verhindern*.

**Datenanalyse** Obwohl immer größere Datenmengen, oft in Form von *Open Data*, öffentlich und frei zugänglich sind und auch die zur Nutzung nötigen Werkzeuge oft kostenfrei zur Verfügung stehen, findet eine Nutzung dieser Daten kaum statt. Datenanalysen bieten die Möglichkeit, durch Kombination verschiedener Datenquellen umfangreiche Informationen zu gewinnen. Typische Methoden wie Clusterbildung, Kategorisierung und Assoziation, können daher in Entscheidungsfindungsprozessen eingesetzt werden. Eine weitere Schlüsselkompetenz ist es daher, zur Verfügung stehende *Daten zu finden, zu nutzen und zu kombinieren, um neue Informationen zu erhalten*. Bei der Durchführung eigener Datenanalysen können die Lernenden nebenbei mögliche Gefahren für die eigene Privatsphäre erkennen: Durch Kombination vielfältiger Datenquellen können z. B. personenbezogene Informationen gewonnen werden, selbst die Deanonymisierung von anonymisierten oder pseudonymisierten Daten ist oft relativ einfach möglich. Die Lernenden sollen deshalb in der Lage sein, *Gefahren für den Datenschutz und die Privatsphäre zu erkennen und bei Datenanalysen ethische Gesichtspunkte mit einzubeziehen*.

**Datenschutz und Datenspuren** Diese Gefahren verstärken sich bei der Nutzung moderner Geräte, Anwendungen und Dienste, die häufig kontinuierlich Daten erfassen. Während

---

die Datenerfassung in einigen Fällen offensichtlich und den Nutzern bewusst ist, z. B. bei sozialen Netzwerken, findet die Erzeugung von Daten oft auch im Hintergrund und ohne Wissen und Zutun des Nutzers statt. Auch Anwendungen bei denen die Datenerfassung offensichtlich ist, erfassen häufig mehr Daten als den Nutzern bewusst ist. Durch die Analyse solcher Daten entsteht die Möglichkeit, große Teile des Privatlebens zu rekonstruieren und für verschiedene – selbst kriminelle – Zwecke einzusetzen. Der Frage ob einer Anwendung vertraut wird, kommt daher eine immer größere Bedeutung zu. Es ist daher eine zentrale Aufgabe der informatischen Bildung, ein Bewusstsein für die Chancen und Gefahren dieser neuen Technologien zu schaffen. Obwohl solche Datensammlungen heute kaum verhindert werden können, kann an vielen Stellen zumindest der Verbreitung der Daten entgegengewirkt werden, indem z. B. Datenschutzeinstellungen genutzt werden. In diesem Zusammenhang stellen sich daher insbesondere drei Fragen: *„Wer speichert welche Daten über mich? Wer kann auf welche Daten über mich zugreifen? Wem kann ich meine Daten anvertrauen?“*. Die in diesem Bereich nötigen Kompetenzen sind daher **eigene Datenspuren zu bemerken** und **mit den Gefahren großer Datenspeicher umzugehen**.

### 3 Fazit

Die beschriebenen Kompetenzen beschäftigen sich mit verschiedenen Aspekten des Umgangs mit Daten und haben einen klaren Bezug zum täglichen Leben. Alle Kompetenzen haben dabei einen zentralen Aspekt gemeinsam: Sie repräsentieren Entscheidungen, die man treffen muss, um Daten angemessen speichern und nutzen zu können. Dies charakterisiert auch die Entwicklungen im Fachgebiet Datenmanagement: Beispielsweise wurden bisher zur Speicherung größerer Datenmengen meist relationale Datenbanksysteme genutzt, während sich heute eine Vielfalt von Datenbanktypen herausbildet, sodass auch hier Entscheidungen, z. B. zur Auswahl des Datenbankmodells, nötig werden. Durch den Erwerb von Kompetenzen auf diesem Gebiet kann gleichzeitig auch den Anforderungen des täglichen Lebens besser begegnet werden: Jeder nutzt heute täglich Anwendungen, die auf Big-Data-Analysen basieren, oft ohne deren Potential und mögliche Konsequenzen und Risiken der Nutzung einschätzen zu können. Gleichzeitig steigt die Bedeutung von Datenmanagement im täglichen Leben stark an: Jeder erzeugt, speichert und verwaltet heute große Datenmengen, muss diese auf eine geeignete Weise speichern und synchronisieren, sowie den Datenschutz und die Datensicherheit sicherstellen. Es kann daher vermutet werden, dass die Relevanz der beschriebenen Kompetenzen in Zukunft noch zunehmen wird, beispielsweise auch durch den Einfluss von Big Data auf vielfältige Berufsgruppen und das Entstehen neuer Berufe, wie z. B. den Datenwissenschaftler („Data Scientist“).

### Literaturverzeichnis

- [GR14a] Grillenberger, Andreas; Romeike, Ralf: A Comparison of the Field Data Management and Its Representation in Secondary CS Curricula. In: Proceedings of WiPSCE 2014. WiPSCE '14, ACM, New York, NY, USA, S. 29–36, 2014.
- [GR14b] Grillenberger, Andreas; Romeike, Ralf: Teaching Data Management: Key Competencies and Opportunities. In: KEYCIT 2014 – Key Competencies in Informatics and ICT. Commentarii informaticae didacticae. Universitätsverlag Potsdam, 2014.