

Tilman Michaeli¹
 Stefan Seegerer¹
 Ralf Romeike¹

¹Didaktik der Informatik, Freie Universität Berlin

Alexa, was ist eigentlich Künstliche Intelligenz?

Einblicke in KI als Gegenstand und Methode fachdidaktischer Forschung

Motivation

Technologische Fortschritte im Kontext der Digitalisierung mit immer leistungsfähigeren Rechensystemen und einer stetig steigenden Datenmenge sind die Grundlage für rapide Entwicklungen im Bereich der künstlichen Intelligenz (KI oder englisch AI). Dabei ist KI kein neues Thema, aber erst im Kontext der Digitalisierung prägt und beeinflusst sie immer mehr Bereiche unseres Lebens. So hat künstliche Intelligenz nicht mehr nur Eingang in Science-Fiction-Filme gefunden, sondern begegnet uns beim Online-Shopping, in Sprachassistenzsystemen, bei ethischen Überlegungen zu autonomem Fahren oder bei maschinellen Übersetzungen. Aber auch in den verschiedensten Fachdisziplinen werden entsprechende Verfahren zur Erkenntnisgewinnung eingesetzt. So werden Datenanalysen und künstliche Intelligenz oft als viertes Standbein der Wissenschaft bezeichnet (Riedel et al., 2008; Wadephul, 2016).

Aufgrund dieser zunehmenden Erfassung, Speicherung und automatisierten Verarbeitung und Auswertung von Daten in allen Bereichen menschlichen Lebens ergeben sich enorme Auswirkungen auf Alltag, Gesellschaft und Politik. Jede und jeder benötigt heute daher Kompetenzen, um die Konsequenzen, Möglichkeiten und Grenzen von künstlicher Intelligenz auf unsere Gesellschaft adäquat und kompetent analysieren, diskutieren und mitgestalten zu können. Die Vermittlung entsprechender Kompetenzen – sogenannter „AI Literacy“ – stellt eine große Herausforderung für Schule, Hochschule und Bildung im Allgemeinen dar. Im Kontext digitaler Bildung wird es zukünftig in allen Fächern verstärkt darum gehen, AI Literacy aufzugreifen. Aus Schülerinnen- und Schülersicht stellen sich hier spannende Fragen, wie etwa, ob Maschinen „denken“ können, wie sie „lernen“ oder welche „Intelligenz“ hinter Phänomene des Alltags, wie beispielsweise Chatbots, Einparkhilfen, Bilderkennung oder Kaufempfehlungen in Online-Shops steckt. Informatische Bildung kann die hierfür notwendigen Grundlagen legen, auf der auch andere Schulfächer fachspezifisch aufbauen.

Dabei gilt es, die Lernenden nachhaltig auf das Leben in der digitalen Welt vorzubereiten. Daher ist es zentral, Kompetenzen zu vermitteln, die – unabhängig von aktuellen Technologien und „Hypes“ – langfristig zum Verstehen und Gestalten dieser digitalen Welt befähigen. Informatik ist jedoch eine vergleichsweise junge und dynamische wissenschaftliche Disziplin und geprägt durch Innovation und sich stetig im Wandel befindliche Technologien und Inhalte. Gerade die fachliche Entwicklung des Forschungsgebiets KI ist nicht abgeschlossen, wie insbesondere die letzten Jahre eindrucksvoll belegen. Vor diesem Hintergrund ist es umso wichtiger, dass fachdidaktische Forschung die zugrunde liegenden Ideen und Prinzipien identifiziert, die unabhängig von konkreten Technologien oder Anwendungen zeitlich überdauernd sind und so zu einer nachhaltigen Vermittlung beitragen. Ein möglicher Forschungsansatz ist die didaktische Explikation nach Grillenberger und Romeike, der in diesem Beitrag aufgegriffen und für das Themengebiet KI adaptiert wird.

Hintergrund

Ansätze wie die Fundamentalen Ideen der Informatik nach Schwill (1993), die Great Principles of Computing (Denning, 2003) oder die Big Ideas of Computer Science nach Bell et al. (2018) strukturieren und charakterisieren die Informatik bzw. ihre Teilgebiete mittels zentraler Begriffe, Ideen, Konzepte oder zugrunde liegender Prinzipien. Durch einen Fokus auf eine zeitüberdauernde Bedeutung, unabhängig von konkreten Technologien, helfen sie das Fachgebiet zu erschließen. Dies erlaubt es, dass „die Bedeutsamkeit eines Themas oder Sachverhaltes überprüft werden kann. Diese Überprüfung liefert wichtige Anhaltspunkte für die Aufbereitung des Stoffes und die Unterrichtsgestaltung“ (Hartmann, Näf und Reichert, 2006).

Allerdings ist festzustellen, dass die Identifikation von zugrunde liegenden Ideen, Konzepten oder Prinzipien der Informatik bzw. eines ausgewählten Teilgebiets wie der künstlichen Intelligenz oft durch epistemologische Überzeugungen oder eingenommenen Perspektiven der Forscherinnen bzw. Forscher beeinflusst wird, sodass eine gewisse Subjektivität nicht ausgeschlossen werden kann (Romeike, 2017; Baumann, 1998).

Um der Normativität solcher Modelle zu begegnen, existieren unterschiedliche Ansätze. Schwill (1993) beispielsweise definiert zunächst eine Reihe von Kriterien, die eine Idee erfüllen muss, um als fundamental für die Informatik zu gelten. Daneben werden etwa Experten aus dem Fach (Zendler und Spannagel, 2006; Zendler, Spannagel und Klaut, 2007) oder Bildungsbereich (Bell et al., 2018) befragt oder aber entsprechende Dokumente wie Lehrbücher oder wissenschaftliche Publikationen analysiert (Armstrong, 2006). Grillenberger und Romeike (2017) entwickelten ein Verfahren zur Strukturierung eines Teilgebiets der Informatik, mit dem in der Analyse von schriftlichen Quellen eine möglichst hohe Objektivität sowie Nachvollziehbarkeit erreicht werden soll, indem ein transparenter Prozess mit kriteriengeleiteten Entscheidungen angewendet wird. Dieses Verfahren der *didaktischen Explikation* besteht dabei aus zwei Phasen (vgl. Abbildung 1).

Im ersten Schritt wird das Fachgebiet explorativ analysiert. Dazu wird zunächst zentrale fachwissenschaftliche Literatur identifiziert, bevor dieses Korpus mittels einer zusammenfassenden qualitativen Inhaltsanalyse untersucht und potentielle Schlüsselkonzepte des Fachgebiets extrahiert und geclustert werden. Die so ermittelten Begriffe werden abschließend mittels einer teilautomatisierten Textanalyse auf Vollständigkeit überprüft. In einer zweiten Phase werden dann die erhaltenen Begriffe kriterienbasiert in einem Modell kondensiert und mit Hilfe von Experteninterviews abgesichert.

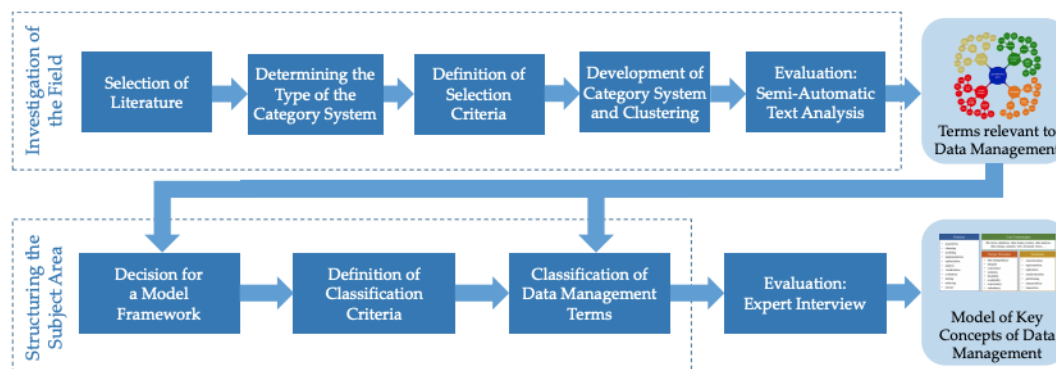


Abb. 1: Didaktische Explikation zur Analyse der Schlüsselkonzepte des Data Management (Grillenberger & Romeike, 2017)

Aufgrund der insbesondere in Phase 1 empirisch geprägten Methodik sind durch die *didaktische Explikation* im Vergleich zu anderen bereits beschriebenen Ansätzen somit Vorteile bezüglich der Güte der Ergebnisse zu erwarten. Nichtsdestotrotz ist das Vorgehen mit einer aufwändigen manuellen Auswertung eines umfangreichen Korpus verbunden und bedarf insbesondere in den Schritten der Extraktion und des Clustering der Schlüsselkonzepte umfangreicher subjektiver Interpretation. Im Folgenden soll daher eine Adaption der *didaktischen Explikation* beschrieben werden, die einerseits den Grad der Automatisierung erhöht, andererseits aber auch zu einer gestiegenen Objektivität und Reliabilität der Vorgehensweise beitragen soll. Skizziert wird die Idee am Beispiel des (dem Datenmanagement thematisch nahen) Fachgebiets der künstlichen Intelligenz selbst: KI analysiert das Fachgebiet KI.

Adaption der didaktischen Explikation

Wie eingangs dargelegt, ermöglichen Verfahrensweisen der künstlichen Intelligenz neue Methoden der wissenschaftlichen Erkenntnisgewinnung und werden daher zur Exploration und Analyse großer Datensätze angewandt. Auch für die *didaktische Explikation* erscheint der Einsatz von Methoden des maschinellen Lernens geeignet, um zu einer weiteren Automatisierung sowie zur Güte des Verfahrens beizutragen. Aufgrund ihres empirischen Charakters bietet insbesondere die erste Phase des Vorgehens hierfür Potential.

Der Einsatz maschineller Lernverfahren und die damit einhergehende Automatisierung ermöglicht es dabei, ein deutlich umfangreicheres Korpus für die Analyse heranzuziehen, was eine höhere externe Validität der Untersuchung gewährleistet – und in der didaktischen Explikation nach Grillenberger und Romeike erst abschließend zur Validierung der potentiellen Schlüsselkonzepte Anwendung findet. Daher wird das Korpus neben entsprechenden Fach- und Lehrbüchern auch um Internetquellen erweitert. Hierzu lassen sich angepasste Web Scraper einsetzen, die neben Texten aus Büchern und wissenschaftlichen Publikationen auch Zeitungstexte, Blog-Artikel oder Tutorials erfassen. Die großen Fortschritte bei Übersetzungsmodellen erlauben es zudem, Quellen aus unterschiedlichen Sprachen zu berücksichtigen und vor Weiterverarbeitung in eine gemeinsame Zielsprache (Englisch) zu übersetzen.

Auf Basis dieses Korpus lassen sich potentielle Schlüsselkonzepte extrahieren. Dazu eignen sich insbesondere Verfahren zur Informationsextraktion wie *Named Entity Recognition*, die es ermöglichen, automatisiert Eigennamen zu identifizieren. Ein gleichzeitiger Abgleich mit Wissensgraphen oder Ontologien zu KI wie DBPedia (Lehmann et al., 2015) erlaubt außerdem die Filterung, Namensauflösung und Disambiguierung dieser Begriffe. Diese Daten können nun aggregiert bzw. hinsichtlich ihrer Häufigkeit und Bedeutung ausgewertet werden. Damit können potentielle Schlüsselkandidaten automatisiert sowie reliabel und objektiv identifiziert werden.

Weiterhin ermöglichen Verfahren des maschinellen Lernens auch, Verhältnis und Vernetzung der identifizierten Begriffe festzustellen und damit das auf die Extraktion folgende Clustering zu unterstützen. So können mit Hilfe von *Topic Modelling* verschiedene Themengebiete der KI in den Dokumenten des Korpus auf Basis einzelner Paragraphen ausfindig gemacht werden, die somit ein inhaltliches Clustering ermöglichen. Darüber hinaus können eigene Modelle für Nomen und Verben trainiert werden, um so beispielsweise eine Unterteilung in *practices* und *principles* (vgl. Denning, 2003) vorzunehmen. Der Einsatz von *Word Embeddings* erlaubt es darüber hinaus, ausgehend vom Kontext der Verwendung, die semantische Ähnlichkeit und die Beziehungen von Begriffen zu bestimmen. Damit können Objektivität und Reliabilität der Interpretation im Kontext des Clustering gestärkt werden.

Somit stellen die solcherart identifizierten und geclusterten potentiellen Schlüsselkonzepte die Basis für Phase 2 der *didaktischen Explikation* dar, in der sie durch die kriterienbasierte Überführung in ein Modell weiter strukturiert werden. Dabei ist aufgrund der dargelegten Adaptionen eine höhere Güte, hauptsächlich hinsichtlich Reliabilität und Objektivität der Ergebnisse, zu erwarten. Nichtsdestotrotz zeigt sich in der Durchführung der Analyse, dass keine vollständige Automatisierung des Verfahrens möglich, sondern immer noch Fach- und Domänenwissen zur Interpretation und Verarbeitung der (Zwischen-)Ergebnisse des Verfahrens notwendig ist.

Fazit

Zusammenfassend zeigt dieser Beitrag damit die Bedeutung künstliche Intelligenz nicht nur als Inhalt digitaler Bildung sondern auch als Methode zur Erkenntnisgewinnung in der fachdidaktischen Forschung auf. So ermöglicht der Einsatz von Verfahrensweisen des maschinellen Lernens die Adaption der *didaktischen Explikation* hinsichtlich einer weiterführenden Automatisierung und Steigerung der Güte des Vorgehens. Verfahren zur Sprachverarbeitung, die auch in persönlichen Assistenzsystemen wie „Alexa“ eingesetzt werden, können also zur fachlichen Klärung als inhaltlicher Grundlage für eine nachhaltige Vermittlung von AI Literacy herangezogen und auch auf andere Teilgebiete der Informatik übertragen werden.

Literaturverzeichnis:

Armstrong, Deborah J. (2006). The Quarks of Object-oriented Development. In: Commun. ACM 49 (2), 123–128.

Baumann, Rüdiger (1998). Fundamentale Ideen der Informatik – gibt es das? In: Informatische Bildung in Deutschland. Perspektiven für das 21. Jahrhundert. Hrsg. von Bernhard Koerber und Ingo-Rüdiger Peters. LOG IN Verlag, 89–107.

Bell, T., Tymann, P., & Yehudai, A. (2018). The big ideas in computer science for K-12 curricula. Bulletin of EATCS, 1(124).

Denning, P. J. (2003). Great principles of computing. In: Communications of the ACM, 46(11), ACM: New York, NY, USA, 15-20.

Grillenberger, A., & Romeike, R. (2017). Key concepts of data management: an empirical approach. In: Proceedings of the 17th Koli Calling International Conference on Computing Education Research, ACM: New York, NY, USA, 30-39.

Hartmann, W., Näf, M. & Reichert, R. (2007). Informatikunterricht planen und durchführen. Berlin, Heidelberg: Springer-Verlag.

Lehmann, J. et al. (2015). DBpedia—a large-scale, multilingual knowledge base extracted from Wikipedia. In: Semantic web, 6(2), 167-195.

Riedel, M. et al. (2008). Classification of different approaches for e-science applications in next generation computing infrastructures. In: eScience'08. IEEE Fourth International Conference on eScience, 2008. IEEE: Indianapolis, IN, USA, 198–205.

Schwill, A. (1993). Fundamentale Ideen der Informatik. Zentralblatt für Didaktik der Mathematik, 25(1), 20-31.

Romeike, R. (2017). Wie informatische Bildung hilft, die digitale Gesellschaft zu verstehen und mitzugestalten, In: Software takes command. Herausforderungen der „Datafizierung“ für die Medienpädagogik in Theorie und Praxis, München: kopaed, 105-118.

Wadepful, C. (2016). Führt Big Data zur abduktiven Wende in den Wissenschaften. In: Berliner Debatte Initial, 27(4), 35-49.

Zendler, A., Spannagel, C., & Klautt, D. (2007). Zentrale Prozesse im Informatikunterricht: eine empirische Grundlegung. In: Notes on Educational Informatics—Section A: Concepts and Techniques, 3(1), 1-19.

Zendler, A., & Spannagel, C. (2006). Zentrale Konzepte im Informatikunterricht: eine empirische Grundlegung. In: Notes on Educational Informatics—Section A: Concepts and Techniques, 2(1), 1-21.