Artificial Intelligence in Education

Systematic literature reviews by Raia Abu Ahmad, Lisa Artman, Erika Monserrat Angelescu, Shantanu Audichya, Kyra Breidbach, Niko Britt, Marlon Dammann, Nele Daske, Niklas Dettmer, Marko Duda, Eliasz Ganning, Leonie Grafweg, Robin Gratz, Isabel Grauwelman, Yannik Heß, Paula Heupel, Fabienne Kock, Henriette Kohnen, Elena Korovina, Lina Lazik, Jakob Lohkamp, Sönke Lülf, Febryeric Malsom Parantean, Ilona Martynenko, Christian Meißner, Juri Moriße, Jacqueline Näther, Felix Naujoks, Nils Niehaus, Tobi Obeck, Alina Ohnesorge, Pia Tamina Ondreka, Cosima Oprotkowitz, Elisa Palme, Carlos Alfonso Parra Fernandez, Michael Rau, Mara Rehmer, Antonella Rönck, Kristina Sigetova, Anna Sommer, Katharina Trant, Henriette Uhlenbrock, Kamran Vatankhah-Barazandeh, Pia Vorsteher, Dennis Witowski, Lara-Sophie Witt, Frederik Wollatz, Liva Zieba & Qirui Zhu, co-edited by Lisa Titz

Edited by

Tobias Thelen

PICS – Publications of the Institute of Cognitive Science
Universität Osnabrück

PICS – Publications of the Institute of Cognitive Science 2022, Number 04

Artificial Intelligence in Education

Systematic literature reviews by Raia Abu Ahmad, Lisa Artman, Erika Monserrat Angelescu, Shantanu Audichya, Kyra Breidbach, Niko Britt, Marlon Dammann, Nele Daske, Niklas Dettmer, Marko Duda, Eliasz Ganning, Leonie Grafweg, Robin Gratz, Isabel Grauwelman, Yannik Heß, Paula Heupel, Fabienne Kock, Henriette Kohnen, Elena Korovina, Lina Lazik, Jakob Lohkamp, Sönke Lülf, Febryeric Malsom Parantean, Ilona Martynenko, Christian Meißner, Juri Moriße, Jacqueline Näther, Felix Naujoks, Nils Niehaus, Tobi Obeck, Alina Ohnesorge, Pia Tamina Ondreka, Cosima Oprotkowitz, Elisa Palme, Carlos Alfonso Parra Fernandez, Michael Rau, Mara Rehmer, Antonella Rönck, Kristina Sigetova, Anna Sommer, Katharina Trant, Henriette Uhlenbrock, Kamran Vatankhah-Barazandeh, Pia Vorsteher, Dennis Witowski, Lara-Sophie Witt, Frederik Wollatz, Liva Zieba & Qirui Zhu, co-edited by Lisa Titz

Edited by

Tobias Thelen

Tobias Thelen (ed.). 2022. *Artificial Intelligence in Education* (PICS – Publications of the Institute of Cognitive Science 2022, Number 04). Osnabrück: Osnabrück University, Institute for Cognitive Science.

This title can be downloaded at:
https://osnadocs.ub.uni-osnabrueck.de/handle/urn:nbn:de:gbv:700-202111125625

Layout and workflows based on awesome templates and tools by Language Science Press, https://langscipress.org

Cover and concept of design: Ulrike Harbort
Fonts: Libertinus, Arimo, DejaVu Sans Mono
Typesetting software: XƎLaTeX

Published by:
Institute of Cognitive Science
Osnabrück University
49069 Osnabrück
Germany
https://ikw.uni-osnabrueck.de

Storage and cataloguing done by Osnabrück University, Universitätbibliothek.

Contents

Vorwort

Vielen Dank, liebe:r Leser:in, dass Sie sich für unsere Sammlung von Übersichtsartikeln zur Künstlichen Intelligenz in der Bildung interessieren. Vielleicht hat Sie Ihr fachliches Interesse auf das Buch aufmerksam gemacht, vielleicht seine Entstehungsgeschichte als Resultat eines Seminar im Studiengang “Cognitive Science” an der Universität Osnabrück. In beiden Fällen hoffen wir, dass auch der jeweils andere Aspekt für Sie interessant ist.

Das vorliegende Buch ist der dritte Versuch, die Teilnehmenden eines Seminars etwas erarbeiten zu lassen, das nicht als Hausarbeit in der Schublade verschwindet, sondern als Sammlung öffentlich zugänglich gemacht werden kann. Von diesen drei Versuchen ist es der zweite, der zu einem tatsächlich veröffentlichten Ergebnis führt und der erste, bei dem der Zeitplan zumindest einigermaßen eingehalten werden konnte. Die vorherigen Probleme lagen nicht etwa daran, dass die Studierenden die Texte nicht rechtzeitig geliefert hätten oder zu lange für die Überarbeitung benötigt hätten, sondern daran, dass ich als Dozent zunächst lernen musste, wie das Seminarkonzept tatsächlich so umgesetzt werden kann, dass ein vorzeigbares Produkt dabei entsteht. Sehr leicht lassen sich nämlich einzelne Schritte der Nachbearbeitung unterschätzen oder es entstehen Versäumnisse, die während des Seminars leicht hätten vermieden werden, anschließend aber kaum noch nachgeholt werden können. So weiß ich z.B. inzwischen, wie wichtig es ist, die Vollständigkeit von Literaturverweisen von Anfang an einzufordern und so früh wie möglich zu kontrollieren.

Kern dieses Seminarkonzeptes ist die Idee, dass in einem ganz regulären Seminar mit zwei Wochenstunden und vier Leistungspunkten ein Ergebnis wie das vorliegende entstehen kann, ohne dass dazu außergewöhnlich hohe Zusatzanstrengungen nötig wären und ohne dass das vorgesehene Zeitbudget deutlich zu überschritten wird.

Ich bin überzeugt, dass dieser letzte Punkt auch bei diesem Versuch nicht eingehalten wurde, denn die Studierenden haben sich für dieses Seminar und dieses Buch in außerordentlichem Maße engagiert. Als Beleg möchte ich nur die Zahl 4.000 anführen, die mich – entschuldigen Sie die Wortwahl in einem wissenschaftlichen Werk – umgehauen hat. Viertausend Kommentare nämlich sind in der gemeinsamen Reviewing-Phase entstanden, in der die fertigen Texte anderer Gruppen von den Studierenden konstruktiv-kritisch gelesen und kommentiert werden sollten. Das ist gelungen, wie so vieles andere auch in und an diesem Buch, dessen Hintergrund nun etwas systematischer erläutert werden soll.

Die Lehrveranstaltung "Artificial Intelligence in Education" im Wintersemester 2021/2022 richtete sich als Wahlpflichtveranstaltung an Bachelor- und Masterstudierende der internationalen Studiengänge "Cognitive Science". Der Kurs war im Modul "Artificial Intelligence" angesiedelt, für die Master-Studierenden aber auch als "Interdisciplinary Course" anrechenbar. Diese Zusammensetzung klingt homogener als sie ist: Die Studiengänge bestehen aus Modulen sehr unterschiedlicher Disziplinen, die auch methodisch sehr unterschiedlich arbeiten. Von der Neurowissenschaft und Neuropsychologie über die Informatik bis zur Philosophie reicht nicht nur die Bandbreite der Lerninhalte, sondern auch (und sogar noch darüber hinaus) die der Interessen und Schwerpunkte der Studierenden. Zudem haben viele Master-Studierende keinen breiten Cognitive-Science-Hintergrund, sondern einen Bachelor in einer der Einzeldisziplinen, so dass sie ggf. kaum Vorkenntnisse zur KI mitbringen.

Das erste Lernziel des Seminares war es, einen Überblick über wichtige Anwendungsfelder, methodische Herangehensweisen und den aktuellen Forschungsstand zum Einsatz Künstlicher Intelligenz für Lehren, Lernen und Bildung zu bekommen. Eine traditionelle Herangehensweise für ein solches Seminar wäre es, dass der Dozent ein bis zwei Dutzend aus seiner Sicht wichtige und repräsentative Paper auswählt und von Studierenden im Semesterverlauf vorstellen lässt. Dieser Ansatz mag aus Lehrendensicht "qualitätsgesicherter" erscheinen, als die Studierenden selbst von Eigeninteresse getrieben nach Themen suchen und dazugehörige Forschungsbeiträge recherchieren zu lassen. Allerdings beweist auch die Themenauswahl in diesem Band wieder einmal, wie viel dabei auch verloren ginge oder gar nicht erst in den Blick geriete, denn eine ganze Reihe von Beiträgen bewegen sich nicht im Kernbereich der "AI in Education"-Forschung, sondern explorieren Ränder und Schnittmengen zu anderen Cogntive-Science-Themen.

Das Seminar teilte sich in drei Phasen auf:

Phase 1 (Mitte Oktober 2021 - Mitte November 2021): Das Seminar startete mit vier vom Dozenten vorbereiteten Sitzungen, die zum einen eine gemeinsame fachliche Grundlage herstellen sowie die verbreitetesten Begriffe, Anwendungsfelder und Forschungsfragen vorstellen sollten. Zum anderen ging es aber auch darum, Gattungen, Publikationswege und Arbeitsprozesse rund um wissenschaftliche Publikationen vorzustellen und anhand kleiner Beispiele und Übungen einzuführen. Schließlich wurden "Systematic Literature Reviews" als Methode zur Erschließung eines wissenschaftlichen Arbeitsfeldes vorgestellt und das Vorgehen bei der eigenen Erarbeitung eines solchen Reviews diskutiert.

Phase 2 (Mitte November 2021 - Anfang Januar 2022): Die Studierenden haben sich an geeignete Themen herangetastet, indem sie entweder in Gruppen nach einem Thema, oder von einem individuellen Interesse ausgehend nach möglichen Gruppen gesucht haben. Die Themenwahl wurde im Seminar diskutiert und reflektiert und die Gruppen haben mit der Literaturrecherche und dem Schreiben der Reviews begonnen. In dieser Phase gab es verschiedene Hilfe- und Unterstützungsformate, bei denen die Studierenden fachliche, stilistische und technische Fragen der LaTeX-Umsetzung mit dem Dozenten und einer Teilnehmerin diskutieren konnten, die sich bereiterklärt hatte, eine unterstütende Editorinnen-Rolle einzunehmen. Die Texte wurden anhand der LaTeX-Vorlage von "Language Science Press" in einer kollaborativen Online-LaTeX-Umgebung (Overleaf über die Academic Cloud Niedersachsen) geschrieben und waren jeweils auf eine Länge von 10 Seiten (exklusive BibTeX-Referenzen) beschränkt.

Phase 3 (Mitte Januar 2022 - Anfang Februar 2022): In einer zweiwöchigen Kommentierungsphase sollten die Studierenden Texte anderer Gruppen kritisch lesen und konstruktiv hinsichtlich der Aspekte "Validität der Befunde", "Vollständigkeit der Argumentation", "Verständlichkeit", "Wissenschaftlicher Sprachstil" und "Einhaltung des Styleguides" kommentieren. Wie oben schon erwähnt, ist dies in außerordentlich umfangreichem Maße geschehen. Viele Kommentare haben außerdem in sehr motivierender Weise Interesse und Begeisterung an der Arbeit der anderen ausgedrückt und alle Anmerkungen sind, auch wenn sie kritisch waren, konstruktiv und respektvoll ausgefallen. An die Kommentierungsphase schloss sich noch etwas Zeit dafür an, die Änderungen einzuarbeiten, die sich aus den Rückmeldungen der Kommiliton:innen ergeben haben, und die Texte zu finalisieren. Die fertigen Texte mussten dann mit Ende der Vorlesungszeit am 6. Februar 2022 eingereicht werden, was auch tatsächlich in allen Fällen hervorragend geklappt hat (hier mögen sich gestandene Wissenschaftler:innen ein Beispiel nehmen).

Das Wintersemester 2021/2022 stand wie die drei Semester zuvor stark unter dem Eindruck der Maßnahmen zur Eindämmung der Corona-Pandemie. Es war als "Hybridsemester" angekündigt, in dem Präsenzveranstaltungen unter Auflagen möglich waren und wieder zum Regelfall werden sollten. Gleichzeitig sollte die Anwesenheit nicht verpflichtend sein, sondern Wahl-Möglichkeiten eröffnet werden. Überdies war das Seminar auch für Studierende im Online-Master-Programm vorgesehen. Es sollte eine vollwertige asynchrone Online-Teilnahme möglich sein, aber das Seminarkonzept sollte auch die Vorteile von

Präsenzsitzungen ausnutzen. Daher wurden die meisten Präsenzsitzungen als Walk-In-Sessions durchgeführt, die für individuelle Fragen und Gruppen-Arbeit genutzt werden konnten, dabei aber nicht aufgezeichnet oder übertragen wurden, sondern sich nur auf die Kommunikation vor Ort konzentrierten. Für die Online-Teilnehmenden gab es ähnlich gestaltete Online-Walk-In-Sessions und individuelle Terminvereinbarungen für Videokonferenzen. Die vorlesungsartigen Teile von Phase 1 wurden asynchron bereitgestellt oder live gestreamt und aufgezeichnet.

Das Semester fiel in ein Spannungsfeld von starken Wünschen nach mehr Präsenz und Rücknahme von Maßnahmen einerseits und stark steigenden Fallzahlen und Diskussionen um Begriffe wie “3G”, “2G” oder “2G+” andererseits. Insgesamt fanden weniger Präsenzsitzungen statt als geplant, das Grundkonzept konnte aber beibehalten werden. Häufige Wechsel und Unsicherheiten bei der Durchführung von Lehrveranstaltungen haben aber insgesamt sicherlich nicht zu erhöhten und längeren Konzentrations- und Arbeitsphasen bei den Studierenden geführt.

Umso erstaunlicher ist es, dass das vorliegende Buch überhaupt entstehen konnte. Wie auch das vorher erarbeitete, aber zeitgleich erscheinende Buch “Artificial Intelligence in Public Discourse” ist es keine Auswahl ganz besonders herausragender Arbeiten oder deshalb zustandegekommen, weil in einem einzelnen Seminar glückliche Umständen zu einem herausragenden Ergebnis führten. Nein, es zeigt, was unsere Studierenden in einer ganz normalen Lehrveranstaltung leisten können und leisten wollen, wenn man ihnen die Möglichkeit gibt, ihre eigenen Interessen und Fragestellungen einzubringen und die ganz offensichtlich motivierende Aussicht eröffnet, etwas Sichtbares zu schaffen.

Nicht alles hat dabei reibungslos geklappt und nicht für alle Teilnehmenden war das Seminar ein Erfolg. Ursache waren zumeist die bekannten Schwierigkeiten, die bei Gruppenarbeit auftreten können. Nicht für alle Studierenden ist es leicht, eine Gruppe zu finden und nicht immer funktionieren Absprachen und Arbeitsteilung in den Gruppen reibungslos. Die Möglichkeiten, solche Konflikte und Probleme zu erkennen und zu lösen, sind in Online-Kommunikations-Szenarien noch einmal eingeschränkter und bei einer erneuten Durchführung des Konzeptes sollte auf die Vermeidung solcher Schwierigkeiten deutlich mehr Wert gelegt werden.

Letztendlich sind 18 Beiträge entstanden, die den Qualitätsanforderungen genügen und deren Autor:innen nach Bekanntgabe der Noten und eines kritischen Feedbacks des Dozenten einer Veröffentlichung zugestimmt haben. Die Gruppen hatten anschließend noch die Möglichkeit, Korrekturen vorzunehmen.

Dieses Buch hat nicht den Anspruch, perfekte Beiträge zu präsentieren. Es ist unter sehr strikten zeitlichen Vorgaben entstanden und dokumentiert auch Lernprozesse einer sehr heterogenen Studierendengruppe: Einige haben das erste Mal im Studium einen wissenschaftlichen Text verfasst, andere hatten vorher kaum Vorkenntnisse zu Künstlicher Intelligenz oder der Anwendungsdomäne "Bildung". Allen gemeinsam ist das große Engagement und der große Einsatz für ihr selbst gewähltes Thema, aber auch für das gemeinsam entstandene Werk.

Die hier nicht enthaltenen gegenseitigen Kommentare zeigen sehr eindrucksvoll, wie konstruktiv, frei von Konkurrenzdenken und mit echtem Interesse an den Themen und Erfolgen der anderen Gruppen die Seminarbeteiligten gearbeitet haben. Im Seminarverlauf ist zu beobachten gewesen, dass die meisten Teilnehmenden zunächst unsicher waren, wie sie ein Thema finden sollten, ob eine Themenidee überhaupt geeignet ist, ob die ersten Recherche- und Zusammenfassungsansätze tatsächlich "richtig" und zielführend sind und schließlich, ob die Texte überhaupt rechtzeitig fertig werden könnten. Zunehmend hat sich dann aber die Identifikation mit dem eigenen Thema und den eigenen Ergebnisse verstärkt und die beobachteten großen Anstrengungen zum Seminarende befördert.

Das vorliegende Buch ist zuallererst und zu guter Letzt allein den Studierenden zu verdanken. Das, was an den Beiträgen gelungen ist, ist den Autor:innen zuzuschreiben, was noch unperfekt und verbesserungsfähig geblieben ist, ist den strikten Vorgaben und meiner kaum ausreichenden Betreuung zuzuschreiben. Während des Seminars, beim Lesen der Beiträge und beim Zusammenstellen dieser Sammlung war ich immer wieder tief beeindruckt, wie sehr die Teilnehmenden das Thema, das Seminar und das Buch zu "ihrem" gemacht haben und bereit waren, ihre Kreativität und Energie einzubringen.

Vielen Dank an alle, die zu diesem Buch beigetragen haben, es war mir eine große Freude mit Euch zusammenuarbeiten!

Ein besonderer Dank geht an Lisa Titz, die als Editorin allen Gruppen hilfreich zur Seite stand und ein wachsames Auge darauf hatte, dass die Beiträge sich an den Style-Guide halten und dass alle Kapitel ausreichend viele Kommentare erhalten.

Nicht zuletzt danke ich dem Verein der Freunde und Förderer des Instituts für Kognitionswissenschaft e.V., der er es möglich gemacht hat, nicht nur eine PDF-Fassung zu produzieren, sondern diese Sammlung auch als echtes Buch zu

drucken. Es macht einen großen Unterschied, die Ergebnisse der eigenen Arbeit auch tatsächlich in den Händen zu halten.

Tobias Thelen, Osnabrück im Juni 2022

Preface

Thank you, dear reader, for your interest in our collection of review articles on Artificial Intelligence in Education. Perhaps your professional interest has drawn your attention to the book, perhaps its origin as a result of a seminar in the course "Cognitive Science" at Osnabrück University. In both cases, we hope that the other aspect is also of interest to you.

This book is the third attempt to let the participants of a seminar work on something that does not disappear in a drawer as a term paper but can be made publicly available as an edited volume. Of these three attempts, it is the second that leads to a published result and the first that at least roughly met the schedule. The previous problems were not because students had not delivered their texts on time or had taken too long to revise them but because I, as a lecturer, first had to learn how the seminar concept could actually be implemented in such a way that it yields a presentable product. It is too easy to underestimate individual steps in the post-processing or to miss out on things that could easily have been avoided during the seminar but can hardly be made up for afterward. For example, I now know how important it is to demand the completeness of literature references from the beginning and to check them as early as possible.

The core of this seminar concept is the idea that a regular seminar with two contact hours per week and four credit points can produce a result like the one presented here, without the need for extraordinarily high additional efforts would be necessary and without significantly exceeding the time budget.

I am convinced this last point was finally not met in this attempt because the students committed themselves to this seminar and this book to an extraordinary degree. As evidence, I would like to cite only the number 4,000, which blew me away – excuse the choice of words in scientific work. The seminar participants contributed four thousand comments during the joint reviewing phase while reading and commenting on the finished texts of other groups. The commenty were supposed to critically and constructively review all papers, and this goal was entirely successful, as was so much else in and about this book, the background to which I will explain somewhat more systematically in the following.

The course “Artificial Intelligence in Education” in the winter semester of 2021/2022 was aimed as an elective course for bachelor’s and master’s students of the international “Cognitive Science” study programs. The seminar was part of the module “Artificial Intelligence”, but for the master’s students, it was also offered as an “Interdisciplinary Course” for credit. This arrangement sounds more homogeneous than it is: The programs consist of modules from very different disciplines, which vary widely in methodology. Topis and students’ interests range from neuroscience and neuropsychology to computer science, philosophy, and even beyond. In addition, many master’s students do not have broad cognitive science backgrounds but a bachelor’s degree in one of the individual disciplines so they may have little prior knowledge of AI.

The first learning goal of the seminar was to get an overview of prominent fields of application, methodological approaches, and the current state of research on the use of artificial intelligence for teaching, learning, and education. A traditional method for such a seminar would be for the instructor to select one or two dozen papers that they consider important and representative and have students present them for the semester. This approach may seem more “quality assured” from the instructor’s point of view than having students themselves search for topics and research papers, driven by self-interest. However, the selection of topics in this volume once again proves how much would be lost or would not even come into view because quite a few contributions do not address core areas of “AI in Education” research, but explore more remote topics and intersections to other cognitive science topics.

The seminar was divided into three phases:

Phase 1 (mid-October 2021 - mid-November 2021): The seminar started with four sessions prepared by the instructor, which were intended, on the one hand, to establish a common understanding and to introduce the most common terms, fields of application, and research questions. On the other hand, the aim was to introduce genres, publication channels, and work processes for academic publishing by employing small examples and exercises. Finally, “Systematic Literature Reviews” were introduced as a method for exploring a scientific field of work. At the end of this phase, we discussed how the participants could conduct their own literature reviews.

Phase 2 (mid-November 2021 - early January 2022): Students approached suitable topics by either working in groups to find a subfield for their review or from an individual interest to find possible groups. The choice of subfields was discussed and reflected upon in the seminar, and the groups began literature searches and review writing. In this phase, the participants could use various help and support formats to discuss professional, stylistic, and technical issues of LaTeXimplementation with the instructor and a participant who had agreed to take a supporting editor role. Texts were to be written using the LaTeXtemplate from "Language Science Press" in a collaborative online LaTeXenvironment (Overleaf via Academic Cloud Niedersachsen) and were limited to a length of 10 pages each (excluding BibTeX references).

Phase 3 (mid-January 2022 - early February 2022): In a two-week commenting phase, students were asked to read critically and comment on texts from other groups concerning the aspects "validity of findings," "completeness of argument," "comprehensibility," "scientific language style," and "compliance with the style guide." As noted above, the participants have been extraordinarily active in commenting. Many comments expressed interest and enthusiasm in each other's work in a very motivating way, and all comments, even if critical, were constructive and respectful. The commenting phase was followed by some time to incorporate the changes that resulted from the fellow students' feedback and finalize the texts. The papers had to be submitted finally at the end of the lecture period on February 6, 2022, which worked out excellently in all cases (here, seasoned scientists may take an example).

The winter semester of 2021/2022, like the previous three semesters, was strongly influenced by the measures to contain the Corona pandemic. The university announced it as a "hybrid semester" in which face-to-face classes were possible under a few restrictions and were to become the norm again. At the same time, classes should not make attendance compulsory but open up options for remote participation. Moreover, the seminar was also intended for students in the online master's program. So I decided that fully asynchronous online participation should be possible, but we should also take advantage of face-to-face sessions. Therefore, most face-to-face sessions were conducted as walk-in sessions that students could use for individual questions and group work. These sessions were not recorded or broadcast and focussed only on on-site communication. For online participants, there were similarly designed online walk-in sessions and individual appointments for videoconferencing. The lecture-style parts of Phase 1 were provided asynchronously or streamed live and recorded.

The semester fell into a tension between solid wishes for more presence and withdrawal of measures on the one hand and enormously increasing case numbers and discussions about terms like "3G", "2G" or "2G+" on the other hand. Overall, fewer face-to-face meetings took place than planned, but we could maintain the basic concept. However, frequent changes and uncertainties in the implementation of classes certainly did not lead to increased and extended periods of concentration and work among the students overall.

Thus, it is all the more astonishing that the present book came into being. Like the book “Artificial Intelligence in Public Discourse”, which was written before but published simultaneously, it is not a selection of outstanding works or came about because fortunate circumstances led to exceptional results. No, it shows what our students can and want to achieve in an ordinary course if they are given the opportunity to bring in their interests and questions and face the motivating prospect of creating something visible.

Not everything went smoothly, and the seminar was not successful for all participants. The cause mostly was the well-known difficulties that can arise in group work. It is difficult for some students to find a group, and group agreements and division of labor do not always function without hassle. The possibilities for recognizing and solving such conflicts and problems are even more limited in online communication scenarios. When I implement the concept again, much more emphasis should be placed on avoiding such difficulties.

In the end, the participants produced 18 contributions that met the quality requirements and for which the authors agreed to publication after they learned about the grade and critical feedback from the lecturer. The groups afterward had the opportunity to make corrections.

This book does not claim to present perfect contributions. It was written under stringent time constraints and documented the learning processes of a very heterogeneous group of students. Some worked on an academic text for the first time in their studies, and others hardly had any previous knowledge about artificial intelligence or the application domain “education”. All groups and students showed a strong commitment and dedication to their self-chosen topic and the work they produced together.

The students’ comments not included here show that the participants worked constructively, free of competitive thinking, and with a genuine interest in the topics and successes of the other groups. In the course of the seminar, it could be observed that most of the participants were initially unsure how to find a topic, whether a topic idea was good at all, whether the first research and summary approaches were actually “correct” and goal-oriented, and finally whether the texts could even be finished in time. However, the identification with one’s topic and results increased and promoted great efforts towards the end of the seminar.

The present book’s success is primarily and solely due to the students. What was successful in the contributions can be attributed to the authors; what remained imperfect and in need of improvement can be attributed to the strict guidelines and my barely adequate supervision. While reading the contributions and compiling this collection during the seminar, I was again and again deeply impressed how much the participants made the topic, the seminar and the book

“their own” and how much they were willing to contribute their creativity and energy.
Big thanks to everyone who contributed to this book. It was a great pleasure working with you!

A special thanks to Lisa Titz, who, as a student co-editor, was helpful to all the groups and kept a watchful eye to ensure the contributions adhered to the style guide and that all chapters received enough comments.

Last but not least, I would like to thank the Verein der Freunde und Förderer des Instituts für Kognitionswissenschaft e.V. for making it possible not only to produce a PDF version, but also to print this collection as a real book. It makes a big difference to hold the results of one’s work in one’s hands.

Tobias Thelen, Osnabrück, June 2022

Introduction

Chapter 1

Introduction

Tobias Thelen

1 What is this book about?

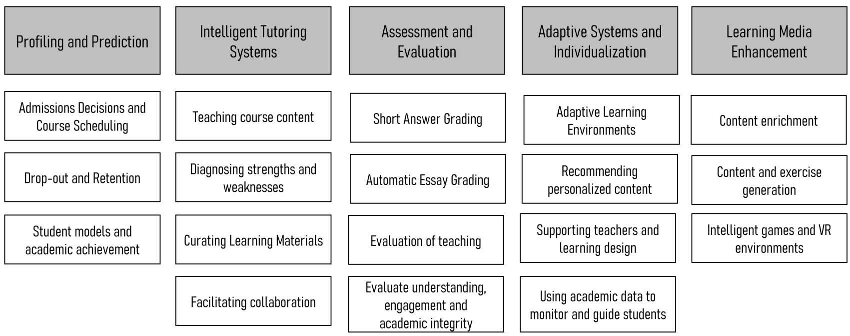


Figure 1: Fields of research and application for AI in education, adapted from Zawacki-Richter et al. (2019)

Educational applications have a very long history in AI research and development. Our society considers education to be a cornerstone of growth and prosperity. At the same time, it is associated with high costs because all citizens should benefit from it comprehensively. It is well-known that individualized education is most effective, but it cannot be automated easily and thus needs a lot of teaching and tutoring personnel.

Thus, using computers in education called for innovative procedures beyond the current state of the art. Graphical user interfaces, processing multimedia data, natural language input, and adaptive systems are connected to early educational computer systems. They tried to decrease the need for human tutors and teachers on the one hand or to increase the quality of education by increasing personalization on the other hand.

The wide range of requirements is reflected in the overview of fields of applications in AI that have been collected at the beginning of the seminar. Figure 1 shows this list of fields. The seminar's purpose was to collect systematic literature reviews summarizing the current state of research, applications, and development in selected subfields of "AI in Education". Thie list of these subfields was not fixed or simply taken from figure 1 but reflected the individual fields of interest chosen by the students.

The papers in this book have been divided into four partially overlapping parts. In the first part, called "Didactics and Ethics", four papers look at paedagogical, didactical, or ethical aspects of AI applications for educational purposes. The "Methods and technologies" part investigates which kinds of AI methods, algorithms, and approaches are currently discussed and which of them show promising results. The fourth part, "Extended reality and robots", includes specialized hardware for using augmented or virtual reality or robots in educational AI applications. The last part returns to the humans involved and their special needs that AI technologies might address in innovative and helpful ways.

2 Didactics and ethics

In this book's first paper, Niko Britt, Marlon Dammann, Nils Niehaus, and Pia Vorsteher look at "A human-centered AIED - the ethical implications of AI in the classroom" (Chapter I). From reviewing current discussions of ethical implications in the educational use of AI, they propose an ethical framework that aims at putting humans – learners and teachers – into focus, not data and algorithms.

The question of unwanted biases as a subfield of ethical discussions attracted much attention and debate in recent years. In their chapter "Biases in AI in education" (Chapter I), Leonie Grafweg, Lara-Sophie Witt, and Paula Heupel give an overview of the types of biases in AI systems, the affected groups of people from an educational application perspective. They summarize the solutions currently discussed in the AI community and point out that the ones affected by educational AI, the young students, are particularly impressionable and in need of protection.

Kyra Breidbach and Antonella Rönck review papers asking about the challenges and opportunities teachers might face when AI technologies are being used in schools. In their article "Consequences of AI for teachers in K-12 education" (Chapter I), they found that a widely discussed field of opportunities is about improved possibilities of providing individualized tutoring and support. At the same time, the technical and conceptual knowledge necessary to realize these opportunities in a paedagogically and ethically responsible way is found to be hard to acquire for teachers.

Picking one aspect of the wide range of consequences for school education, Fabienne Kock and Anna Sommer investigate "AI teacher assistants" (Chapter I). They looked for actual research-backed applications that aim at supporting teachers utilizing supportive chatbots and learning analytics. They conclude that there is a lot of activity in the field, and many applications are being developed, but they are not ready for widespread general use yet.

Pia Tamina Ondreka, Carlos Alfonso Parra Fernandez, Frederik Wollatz, and Qirui Zhu take a specialized look at "AI-based adaptive gamification in math learning scenarios" (Chapter I). As they could not find a sufficiently large body of academic works addressing their question, they started summarizing the findings of gamification and math learning applications separately. In conclusion, they provide evidence that using adaptive gamification for math learning could be beneficial, especially for students with math anxiety.

3 Methods and technologies

Marko Duda and Niklas Dettmer investigate how a particular class of algorithms for presenting learning tasks is discussed in current research. Their paper "Comparing different spaced repetition algorithms for enhancing human learning" (Chapter II) reports long-lasting research efforts in optimizing spaced repetition learning and finds that both rule-based and ruleless algorithms have their respective fields of application.

Individualization and personalization are key aspects of AI applications for education. Shantanu Audichy and Elena Korovina take a look at how traditional machine learning and more recent deep learning approaches are applied in their paper "Personalized learning experiences" (Chapter II). They find recommendation algorithms to be the basis of many personalized learning approaches. They also identified the availability of large data sets as being crucial for application and research.

This topic is addressed in the paper "Data sets in AI in education" by Erika Monserrat Angelescu and Jacqueline Näther (Chapter II). They present how data privacy regulations create the need for anonymization techniques. In their overview, different methods are discussed, and they find the research field to be active and yield promising results already, but the need for large data sets is even more extensive than what has been accomplished so far.

Nele Daske, Raia Abu Ahmed, and Febryeric Malsom Parantean reviewed a specialized application domain, the "Automatic lecture transcription" (Chapter II). The literature survey conducted for this paper aims to show the variety of methodological approaches. They show that choosing technologies for creating the acoustic and language models is essential but can only go together with an appropriate selection of training data and evaluation metrics. The authors find the field of research a bit inactive and believe that more recent NLP approaches could improve the performance of lecture transcriptions.

A common problem for lecture transcription is the recognition of domain-specific vocabulary. On a more fundamental level, Kristina Sigetova, Eliasz Ganning, and Tobi Obeck address the issue of domain-specific aspects in their paper "Knowledge representation approaches in adaptive educational AI systems" (Chapter II). They find recommender systems and ontologies to be the dominant research directions in this field, with recent activities in solving the cold-start problem for newly created systems by using hybrid approaches and other AI methods.

4 XR and robots

Extended reality (VR/AR) is a much-researched field of innovative hardware uses and is often connected to AI algorithms. Jakob Lohkamp, Felix Naujoks, and Juri Moriße investigate literature on "Combining AI and virtual reality in medical Education" (Chapter IV). They find AI being applied to case-based and problem-based learning scenarios that include VR/AR simulations that might have the potential to make medical training provision much easier. From the papers investigated, the authors identify a need for a better understanding of the effects of such simulations, also from ethical and diversity perspectives.

Henriette Kohnen, Alina Ohnesorge, Henriette Uhlenbrock, and Dennis Witowski wanted to know about the current state of research on a topic they called "Making science fiction reality – applications, benefits and drawbacks of 3D holograms in education" (Chapter III). Interactive 3D holograms promise new forms of human-computer interaction but technologically seem to be far away. The authors present an overview of historical and current approaches to making holograms usable and various possible applications in educational settings. They show that the technology is on the threshold of actual usability but is currently associated with very high costs that generally prevent practical use.

A bit further in terms of practical implementation is the topic of Robin Gratz, Lina Lazik, Sönke Lülf, and Elisa Palme and their paper “Humanoid robots in Education” (Chapter III). They can show that researchers see great potential in using robots in the classroom and provide examples from all kinds of formal education. The roles of robots in these scenarios are manifold and range from teachers to tutors and assistants to peers. However, there are ethical questions that must not be neglected.

The last paper in this part could also be the first paper of the next part as it combines both main aspects: “AI-based virtual reality systems for helping learners with special needs” is the title of the review by Yannik Heß and Michael Rau (Chapter III). They investigate “VR scaffolding” to create a specialized and enriched environment to facilitate learning. The authors identify a great potential to make education more accessible for several groups of students with special needs. Still, they also show that more research and development efforts must be put into this field.

5 Addressing special needs

The capability of AI-driven education support systems to create specialized learning environments for students with special needs is also addressed in the paper “AI approaches towards reducing the barriers for learners with sensory or physical impairments” by Kamran Vatankhah-Barazandeh and Christian Meißner (Chapter IV). They give an overview of developments and research from the last twenty years, indicating that AI has indeed the potential to provide better access to education for learners with different kinds of impairments.

Mara Rehmer and Katharina Trant focus on “AI-typical learning buddy – A literature review of AI-based learning applications for people in the autism spectrum” (Chapter IV). They find applications for children to help with verbal communication, emotion detection, and emotion expression. They conclude that current research lacks the inclusion of larger test groups and typically provides too little technical detail information.

One of these aspects, communication support, was investigated closer by Lisa Artmann, Ilona Martynenko, and Liva Zieba in their paper “Using AI for enhancing communications skills and strategies of children with autism spectrum disorder” (Chapter IV). In the research literature taken into account for this study, many technologies encountered before in this book, like augmented reality, robots, and virtual agents, are explored for their potential to help. The authors find that most studies in this growing field have a technological rather than psychological focus and, therefore, might use technology just because it is available. But still, progress in the area appears to be very promising.

The last paper by Cosima Oprotkowitz and Isabel Grauwelman is titled "AI in education for children with intellectual disabilities" (Chapter IV). They find their field of application to be relatively new, and it did not receive much attention from AI in education research yet. However, adaptive learning systems show promising results and might reduce emotional stress because they could reduce negative social aspects in learning groups. The authors point out that AI technology can not compensate for other shortcomings in treating children with intellectual disabilities in the educational system.

6 Summary – What did we learn about AI in education?

Summarizing the papers differently with such a diverse and broad range of questions and subfields is nearly impossible. But it can be stated that even after more than 60 years of AI in education research and recent breakthroughs in, e.g., image processing and natural language processing, AI has not (yet?) fully arrived in education. There are a lot of promising approaches, mostly combining older ideas with more recent methods. However, AI still has a way to go as a paedagogically valuable support tool for teachers and learners.

This critique is in no way meant to minimize the merits in detail. Especially where general-purpose AI technologies fit for education, they are adopted quickly and find practical applications, as became apparent when looking at virtual reality, chatbots, recommender systems, or humanoid robots. As a specialized field of application, education remains a challenging problem for AI. Teaching and learning are profoundly human, and today's technology can hardly achieve the beneficiary effects of an individual human tutor.

References

Zawacki-Richter, Olaf, Victoria I. Marín, Melissa Bond & Franziska Gouverneur. 2019. Systematic review of research on artificial intelligence applications in higher education–where are the educators? *International Journal of Educational Technology in Higher Education* 16. 39. DOI: 10.1186/s41239-019-0171-0.

Part I

Didactics and Ethics

Chapter 2

A human-centered AIED – The ethical implications of AI in the classroom

Niko Britt, Marlon Dammann, Nils Niehaus & Pia Vorsteher

The topic of ethics in artificial intelligence in education (AIEd) is being neglected in many different areas. Currently, as shown by Schiff (2021) and Bozkurt et al. (2021), there is a significant lack of policies and academic papers pertaining to the ethics of AIEd, as well as an urgent need for guidelines to be defined and adopted (Holmes et al. 2021). This review sheds light on the often-ignored notion of ethics in AIEd by conjoining the work of researchers found through a systematic online search. We adopt the ethical principles developed by Floridi & Cowls (2021) and create our own framework by applying them to AIEd. The framework includes the following ethical principles: beneficence, non-maleficence, autonomy, transparency, and justice. We conclude that, to ensure compliance with all mentioned ethical principles, any development in AIEd must protect the rights of everyone involved. Seeing as (data set) bias in AIEd is an important and extensive topic, we further elaborate on the different characteristics data set bias can be based on gender, ethnicity, background, disability status, dialect, and socioeconomic status. To ensure ethical AIEd by, for example, reducing data set bias, we need to diversify the community of people who are working on addressing the concerns AIEd brings about.

Keywords: Ethics, Artificial Intelligence, Education, Privacy, Bias

1 Introduction

- How might students and teachers benefit from AIEd technology?
- "What are the ethical obligations for developers of AIEd technology and researchers in the field?" Holmes et al. (2021: 6)

- "What are the ethical implications of not being able to easily interrogate how some AIEd deep decisions [...] are made?" Holmes et al. (2021: 6)

These and more are ethical questions which will be addressed in this paper. We will attempt to give an overview of the most important ethical principles the AIEd community should abide by to ensure ethical soundness.

AI is on an ongoing path of expansion and is starting to influence the educational sector. With the growing interest AIEd has received in recent years, the number of research papers in the field per year is increasing. However, upon further inspection, one can observe a research gap when it comes to the ethics of AIEd. For this reason, it is important to highlight ethical implications of the future of AIEd and give students and teachers a louder voice. We attempted to do exactly that by consolidating many different papers on the topic found through a methodical online search.

2 Methodology

We conducted our search for papers by employing two search strategies: first, we used two different data bases, namely Scopus and Google Scholar, and then looked at the works that they cited individually. The keywords we chose to search the databases were: Ethics in AI, Ethics in AIEd, Philosophy of AI, Philosophy of AI in education, Philosophy in AIEd, and AIEd data set bias. This produced around 2 million results combined. To make the search more bearable, we filtered the results by relevance, which is based on the citation count among other factors and went through the first 3 pages for each term for both data bases. We then manually filtered out all the works irrelevant to our cause by reading the title and abstract of each of the 260 papers we looked at. The choice of papers was limited, in the sense that we only included accessible papers written in English or German. Our search strategies allowed us to find 25 relevant sources that are mostly journal articles published around 2019. Of these 25 works, 5 were thoroughly read and more heavily cited by us, namely Floridi & Cowls (2021), Holmes et al. (2021), Baker & Hawn (2021), Schiff (2021), and Holstein et al. (2019), while the other 20 were mentioned in our paper to underline a specific matter or give further references to the reader.

In the following sections, we will elaborate on our findings, starting with why the topic of ethics in AIEd is worth considering.

3 Importance of Ethics in AIEd

Even if AIEd researchers are driven by ethically good intentions, this does not suffice for an ethical implementation of AIEd (e.g., Dastin (2018), Reich & Ito (2017), Whittaker et al. (2018)). It must be ensured that AIEd is ethical by design, not merely by intention (Holmes et al. 2021). Unfortunately, as stated by Holmes et al. (2018): "Around the world, virtually no research has been undertaken, no guidelines have been provided, no policies have been developed, and no regulations have been enacted to address the specific ethical issues raised by the use of Artificial Intelligence in Education." Holmes et al. (2018: 552).

It is important to research AIEd in the context of ethics, because there is a significant lack of academic papers pertaining to the ethics in AIEd according to Bozkurt et al. (2021). Also, key national AI policy strategies that discuss the impact of AIEd are neglecting the ethical implications of predictive assessments or intelligent tutoring systems (ITS) and other "traditional technologies" of AIEd as discussed by Schiff (2021). He emphasizes that the examined documents released by over 30 countries arguably illustrate the highest priorities of global policy makers. These policy makers seem to ignore, or at least not attend to, the risks of unethically implemented AIEd developments. Rather, they focus on training AI(Ed) experts and preparing the workforce for AI(Ed), as the analysis of Schiff (2021) concludes. According to him, policy discourse disregards the ethical implications of AIEd, because it rather focuses on (economic) growth and sees AIEd as an instrument to achieve that growth. To further public and policy discourse in AIEd, an awareness of the ethical implications must be established. Only with the help of ethical consideration and appropriate resulting policies can we ensure a responsible implementation and governance of AIEd (Schiff 2021).

A study conducted by Holmes et al. (2021) had 17 of the leading researchers in the field of AIEd fill out a survey relating to ethics in AIEd. The results found that researchers agreed on the need for improvement when it comes to the attention ethics in AIEd is receiving. According to several respondents, AIEd lacks an urgently needed framework to deal with its ethical implications.

4 Our AIEd Ethics Framework

A thoughtfully devised AIEd policy framework with a careful consideration of possible ethical implications has enormous potential for guiding AIEd research and development, benefitting the entire educational sector. For the field of AIEd, we need to account for both AI ethics as well as ethics in education. The AI ethics principles guiding our framework were developed by Floridi & Cowls (2021), additionally, Aiken & Epstein (2000) provided a detailed account of ethical guidelines for both AIEd and general Computer Science which were also considered. Principles from Floridi & Cowls (2021) are non-maleficence, beneficence, autonomy, transparency, and justice. Originally, they were synthesized from 47 ethics principles, implying that most other principles can be mapped to this framework (e.g., safety, fairness, etc.). Similar to Schiff (2021), we are going to apply these principles to AIEd.

4.1 Non-Maleficence and Privacy

Non-Maleficence emphasizes that no harm must be put upon others. This principle encompasses privacy and the manipulation of children, among others.

Privacy is a major concern when it comes to non-maleficence in AIEd. Here, the conflict lies between surveillance exercised by AIEd systems through, for example, large scale data collections or cameras that monitor students, and the students's right to privacy. UNICEF (2020: 30) defines someone's data as any information collected about, content created by, or algorithmic inferences pertaining to them. The harm involved in collecting large quantities of data from students relating to their academic performance and competencies is considerable. Any entity inevitably gains power by collecting data. While a small amount of surveillance has the potential to create a safer and more productive environment, constant student monitoring does not. Rather, it can cultivate social control and self-censorship which infringes on the student's freedom and jeopardizes creativity and self-determination (UNICEF 2020). As a specific example of student monitoring, we can take a closer look at one of the newest features of ITSs: human emotion detection. This technology allows the AIEd system to monitor and measure the student's behavior by looking at their eye movement or other stimuli and inferring their attentiveness, for example. Consequently, broad access to this technology could lead to the student receiving punishment and would possibly target the weakest students such as those with learning disabilities (Schiff 2021). Thus, AIEd technologies such as these must be implemented with caution and consideration, if at all. Privacy must be guaranteed by design, and the students should be the ones who have control over the data through use of informed consent or rejection. For example, a student's analytics should be hidden from the teacher and the system by default, and access should only be granted if the student allows it.

4.2 Beneficence

It must be emphasized that an ethical evaluation of AIEd does not merely aim at preventing researchers and developers from "doing harm", but it also aims at promoting benevolent objectives. Beneficence is often attributed to human values and the social good. It demands that one must contribute to the welfare of others. AIEd can abide by this principle when it is not being solely employed to further economic goals but rather focuses on the purpose of learning. The pedagogical approaches implemented are to be well thought through. The current approach to education, instructionism, meaning a paradigm with an instructor role and a fixed or less individualistic schedule, is contested (Holmes et al. 2021). Student rights and well-being must be the primary focus. The benefit AIEd systems and policies can provide to students must be prioritized. One instance of this could be an AIEd system that contributes to the welfare of learners by ensuring broad access to education thanks to AI's scalable and adaptable nature. It can and should be used in an inclusive manner to grant any student, no matter the socioeconomic background or skill level, an opportunity at quality education. This is especially useful for students living in less-developed countries or for slow learners (Schiff 2021).

4.3 Autonomy

In AIEd, especially when implemented on a large scale, decision-making is often being outsourced to AIEd systems. This can lead to a loss of autonomy on the side of the teachers, students, and parents. The lack of oversight can result in unwanted consequences such as the inability to adapt to it or control it. A recent example from the UK is the faulty algorithm for grade predictions for university admissions, employed by Ofqual (Hao 2020). It follows that, because AIEd is not always accurate, important decisions that have a great impact on the student need to be overseen by a human that makes the final decision.

4.4 Transparency

Without transparency, biases or errors in an algorithm are practically invisible to anyone who does not have the data or access to the algorithm. Students, teachers, and parents dissatisfied with the AIEd systems' interpretation of their performance could not receive sufficient explanations.

AIEd developers and policy makers must be held accountable to prevent, among other things, the exploitation of a child's vulnerability or lack of understanding (UNICEF 2020). A child has the right to know the features and implications of the AI system it interacts with and so do their parents or caregivers. The creation of a mechanism or institution to ensure compensation and justice in the case of any transparency or accountability infringements would be beneficial, according to (UNICEF 2020). If transparency and accountability are not practiced, justice cannot be pursued.

4.5 Justice

UNICEF (2020: 28-30) defined requirements that aim at ensuring justice in AI. According to the authors, diversity amongst those who design, develop, implement, research, regulate and oversee AI systems must be guaranteed. Because bias, being the main component of justice in the context of AIEd, is such a multi-facetted topic, we will dissect the issue in the following paragraphs.

5 Biases in Data sets and Algorithms

Almost every training algorithm for automatic learning is based on the data set that is used for training. If the algorithm design itself is inherently susceptible to bias or the data set is biased, the algorithm and the intelligent device gets influenced accordingly. This influence is referred to as data set or algorithm bias depending on where the limitations lie and can have negative effects that can lead to the discrimination or exclusion of people or certain groups of people.

Data set bias is just one of several biases that currently exist. It seems relevant to mention that there are many biases whose potential existence researchers have yet to provide evidence for. These biases are based on disability status, dialect, socioeconomic status, urbanicity, native language, national region, parental education background, military-connectedness and migrant work (Baker & Hawn 2021).

Certain existing technologies point to the fact that racial and gender biases are not a novelty in the field of AI. For example, Google's automatic sentence completion for the words black and white, or woman and man, have yielded biases results in the past (Zou & Schiebinger 2018) and Amazon's models have had alleged cases of gender bias (Dastin 2018). The generalizations embedded into these technologies are also embedded into technologies used in artificial intelligent systems in schools. Thus, they could cause similar problems.

In the following, we are going to look at three different major categories within bias in AIEd: ethnicity, nationality, and gender more closely.

5.1 Bias: Ethnicity, Nationality, Gender

Some AIEd algorithms favor or disfavor one ethnic group over another. For example, different studies investigated the performance of an algorithm that predicts if a student is at risk of failing a course or not between Black and White students. Hu & Rangwala (2020) found that results were inconsistent across university courses, whereas, in another study, Anderson et al. (2019) showed a higher than average false positive rate for White students and a higher than average false negative rate for Latino students.

Nationality is also a major characteristic algorithmic biases are based on. In an E-rater system, Chinese and Korean students got higher scores on a test for foreign language proficiency than they did from human essay raters (Bridgeman et al. 2009). A similar study found that speakers of Arabic and Hindi were given lower scores (Bridgeman et al. 2012).

Moreover, researchers found that a model trained on data from the United States was highly accurate for students from other economically developed countries but less accurate for students from less economically developed countries (Li et al. 2021).

When investigating gender bias, researchers looked at algorithms for predicting if a student is at risk of failing a course which exhibited a worse performance for male students compared to female students (Hu & Rangwala 2020). However, the results were inconsistent across university courses. Another study concerning the performance between male and female students in a model predicting six-year college graduation across five different algorithms, revealed higher false negative rates for male students than for female students (Anderson et al. 2019). However, courses with a high proportion of male students (50-80 percent), showed worse performance for female students than for male students (Gardner et al. 2019). For both 11th grade essays and foreign-language proficiency examinations, the E-Rater system seemed to be evenly accurate for both male and female students (Bridgeman et al. 2009). Information about algorithmic bias against non-binary or transgender learners was not brought up (Baker & Hawn 2021).

Most of the studies mentioned above seem to provide strong evidence of the existence of biases in AI algorithms integrated in the educational system. They demonstrate the apparent injustice that students have experienced and will continue to experience if nothing is done to change how AIEd implementations are dealt with.

Bias against any person or any group of people must be eradicated. In the following, we will outline how students, teachers, and researchers could go about combating the issue of bias.

5.2 How Researchers, Teachers, and Students can address Bias

According to Baker & Hawn (2021), to fight biases, researchers should collect more data on group membership in the data collection stage. Data on ethnicity, gender, and national origin should be included to allow for an analysis of the data sets for algorithmic biases whenever possible. Baker & Hawn (2021) argue that researchers could also reduce bias by improving the labelling process which occurs during data collection. Labels coming from external sources can themselves be biased. Thus, according to them, it may be useful to make training labels more objective where possible or, as has been suggested by Okur et al. (2018), to have people from affected groups do the labelling to ensure less bias. Baker & Hawn (2021) argue that the developers' gender and background distribution can also play a role and propose that this should be diversified to avoid biases. The authors conclude that AIEd researchers should critically reflect on their own research and developments.

Teachers could try to join interdisciplinary teams addressing bias in AIEd. They may need specific training courses to be taught in how to supervise AIEd activities and how to integrate them with others (Qian 2021).

Students are not aware of many biases (Qian 2021). Therefore, ethics researchers agree that in order for students to combat bias, they must first be informed of its existence and implications. For example, Melsión et al. (2021) argue that children need to be aware of issues such as transparency and fairness so that they can identify the societal impacts of bias in AI. Therefore, their study leveraged state of the art explainability techniques to help children better understand gender bias in AI. Specifically, they used a visualization tool to direct attention to the presence of gender bias in certain research data sets with the objective of making explicit that the traditional view of gender, here, the binary gender construct and sexist stereotypes, exists in the data collected to train certain machine learning models. The bias visualization tool helps students do three things: it helps recognize that a machine learning system might perform biased predictions, it aids in identifying biased models, and it helps students come to the realization that training data plays an essential role in the production of bias (Melsión et al. 2021).

6 What do Teachers and Students want from AIEd?

Since students and teachers are often not aware of the more subtle dangers of AIEd technology, the designing parties have a moral obligation to be fair and transparent. A good starting point when faced with AIEd system design decisions, is to ask the impacted parties directly what they want and what they do not want in an AIEd system. However, as discussed before, teachers and students can only give adequate answers if they are informed.

Employing Explainable AI (XAI) which involves informing the users on how and why a decision was made and what that decision entails, is not only helpful for education directly but is also in compliance with the ethical principles of transparency and accountability. This in turn can create trust in the technology among users (Putnam & Conati 2019).

Despite the evidence suggesting that explanations are useful and effective in general, other studies find that this varies between systems. Putnam & Conati (2019) found that while users show an overall positive attitude towards explainable recommender systems, low-cost intelligent interactive systems like the "Google" search algorithm seem to be better off, as users were uninterested or in some cases even annoyed by it. With their study the authors tried to gain a better understanding of students' opinions on the matter of AI applications and the use of explainable AI. A group of nine students participated in an experiment that involved working with an ITS that solves constraint satisfaction problems. The participants reported on the type and usefulness of the given explanations. The results support what was described earlier. While results vary between types of tasks and situations, most students wanted more explanations and transparency.

Nazaretsky et al. (2021) outlines a more concerning approach to AI systems which aims at supporting teachers and maybe eventually replacing them completely. The research experiment Nazaretsky et al. (2021) refers to consists of 16 science teachers that participate in a yearlong training program. In this program, teachers design and simulate features for an AI-based learning environment, in this case the free environment "PeTeL", in order to grant the AI the ability to analyze and grade students' data as well as provide further recommendations.

According to Nazaretsky et al. (2021), aside from noticeable improvements in the efficiency and correctness of the evaluations, the use of such tools leads to the emergence of issues among teachers. Besides the fact that teachers are forced to change their teaching habits, a main issue is the presence of a confirmation bias towards the teacher's own perception of the learner, over the AI's perception. Teachers noticed that AI results were incongruous with the image they had of certain students, especially with those students they knew better and longer. This further emphasizes the importance of teacher-student relations. It was also evidenced by certain implications and expressions only the teachers were able to understand.

A promising example that serves as a good ethical basis for some aspects of AIEd design is provided in Holstein et al. (2019). The foundations of these findings are the result of a participatory speed dating (PSD) session with students and teachers in which both parties discussed presented designs involving ITSs Holstein et al. (2019: 1). Following the discussions, a concept approval rating was created and the results were discussed in detail in Holstein et al. (2019: 5-10).

Most preferred concepts by teachers were ”feedback on teacher explanations” and ”student rankings for receiving help”. Least preferred was ”granting full privacy to students to let them hide important analytics from the teacher completely”. For students, the most preferred concept was for the system to ask the student for permission before sharing specific, often emotionally associated, analytics with the teacher. Least preferred were concepts in which parents or student peers would gain too much insight into their performance analytics, as shown in Holstein et al. (2019: 5-10).

7 Conclusion

So far, neither the student nor their education is at the core of AIEd. Currently, AIEd primarily concentrates on data and computation. Ethical topics in AIEd are receiving little attention. Ethics in AIEd needs to be taken seriously. Therefore, we defined a much-needed ethical framework that everyone in the AIEd community can abide by. In doing so, the AIEd community can assist in ensuring that AIEd systems are ethically implemented, especially those affecting impressionable young people. Informing teachers and students about issues as well as possibilities of AIEd technology is important to ensure that their needs can be met. Students had an overall positive and teachers a slightly more negative stance towards AIEd. However, both groups had a mostly positive attitude towards an explainable AI. The fact that student-teacher relations and in-person education are important and, in some cases, even more appropriate than the forms of learning AIEd technology can provide, must not be overlooked.

The ethical issues listed in this paper make up only some of the known ethical implications AIEd can have. Also, AIEd will bring with it unknown implications that have yet to be discovered. The AIEd community, students, teachers, policymakers, and stakeholders need to work together to bring these to light and establish and implement an extensive framework for an ethical future of AIEd.

References

Aiken, Robert M. & Richard G. Epstein. 2000. Ethical guidelines for ai in education: starting a conversation. *International Journal of Artificial Intelligence in Education* 11. 163–176.

Anderson, Henry, Afshan Boodhwani & Ryan S. Baker. 2019. Assessing the fairness of graduation predictions. In *12th international conference on educational data mining (edm 2019)*, 488–491.

Baker, Ryan Shaun & Aaron Hawn. 2021. Algorithmic bias in education. *International Journal of Artificial Intelligence in Education* 31(4). 1–41. DOI: 10.1007/s40593-021-00285-9.

Bozkurt, Aras, Abdulkadir Karadeniz, David Baneres, Ana Elena Guerrero-Roldán & M. Elena Rodríguez. 2021. Artificial intelligence and reflections from educational landscape: a review of ai studies in half a century. *Sustainability* 13(2). 800. DOI: 10.3390/su13020800.

Bridgeman, Brent, Catherine Trapani & Yigal Attali. 2009. Considering fairness and validity in evaluating automated scoring. In *Annual meeting of the national council on measurement in education (san diego, ca, usa). ncme, mt. royal, nj*, 1–18.

Bridgeman, Brent, Catherine Trapani & Yigal Attali. 2012. Comparison of human and machine scoring of essays: differences by gender, ethnicity, and country. *Applied Measurement in Education* 25(1). 27–40. DOI: 10.1080/08957347.2012.635502.

Dastin, Jeffrey. 2018. Amazon scraps secret ai recruiting tool that showed bias against women.

Floridi, Luciano & Josh Cowls. 2021. A unified framework of five principles for ai in society. *Harvard Data Science Review*. DOI: 10.1162/99608f92.8cd550d1.

Gardner, Josh, Christopher Brooks & Ryan Baker. 2019. Evaluating the fairness of predictive student models through slicing analysis. In *Proceedings of the 9th international conference on learning analytics & knowledge*, 225–234. New York: Association for Computing Machinery. DOI: 10.1145/3303772.3303791.

Hao, K. 2020. *The uk exam debacle reminds us that algorithms can't fix broken systems.* https://www.technologyreview.com/2020/08/20/1007502/uk-exam-%20algorithm-cant-fix-broken-system/.

Holmes, Wayne, Duygu Bektik, Beverly Woolf & Rose Luckin. 2018. Ethics in aied: who cares? In *Artificial intelligence in education: 19th international conference, aied 2018, london, uk, june 27–30, 2018, proceedings, part ii*, vol. 10948, 551–553. Cham. DOI: 10.1007/978-3-319-93846-2.

Holmes, Wayne, Kaska Porayska-Pomsta, Ken Holstein, Emma Sutherland, Toby Baker, Simon Buckingham Shum, Olga C Santos, Mercedes T Rodrigo, Mutlu Cukurova, Ig Ibert Bittencourt, et al. 2021. Ethics of ai in education: towards a community-wide framework. *International Journal of Artificial Intelligence in Education*. 1–23. DOI: 10.1007/s40593-021-00239-1.

Holstein, Kenneth, Bruce M. McLaren & Vincent Aleven. 2019. Designing for complementarity: teacher and student needs for orchestration support in ai-enhanced classrooms. In *International conference on artificial intelligence in education*, 157–171. Cham: Springer. DOI: 10.1007/978-3-030-23204-7_14.

Hu, Qian & Huzefa Rangwala. 2020. Towards fair educational data mining: a case study on detecting at-risk students. In *13th international conference on educational data mining (edm 2020)*.

Li, Xingyu, Difan Song, Miaozhe Han, Yu Zhang & Rene F. Kizilcec. 2021. On the limits of algorithmic prediction across the globe. *arXiv preprint arXiv:2103.15212*.

Melsión, Gaspar Isaac, Ilaria Torre, Eva Vidal & Iolanda Leite. 2021. Using explainability to help children understandgender bias in ai. In *Interaction design and children*, 87–99. New York: Association for Computing Machinery. DOI: 10.1145/3459990.3460719.

Nazaretsky, Tanya, Mutlu Cukurova, Moriah Ariely & Giora Alexandron. 2021. Confirmation bias and trust: human factors that influence teachers' attitudes towards ai-based educational technology. In *Ceur workshop proceedings*, vol. 3042.

Okur, Eda, Sinem Aslan, Nese Alyuz, Asli Arslan Esme & Ryan S. Baker. 2018. Role of socio-cultural differences in labeling students' affective states. In *International conference on artificial intelligence in education*, 367–380. Cham: Springer.

Putnam, Vanessa & Cristina Conati. 2019. Exploring the need for explainable artificial intelligence (xai) in intelligent tutoring systems (its). In *Joint proceedings of the acm iui 2019 workshops co-located with the 24th acm conference on intelligent user interfaces (acm iui 2019)*.

Qian, Zhang. 2021. Applications, risks and countermeasures of artificial intelligence in education. In *2nd international conference on artificial intelligence and education (icaie), 2021*, 89–92. New York: IEEE. DOI: 10.1109/ICAIE53562.2021.00026.

Reich, Justin & Mizuko Ito. 2017. From good intentions to real outcomes: equity by design in learning technologies. *Digital Media and Learning Research Hub*.

Schiff, Daniel. 2021. Education for ai, not ai for education: the role of education and ethics in national ai policy strategies. *International Journal of Artificial Intelligence in Education*. 1–37. DOI: 10.1007/s40593-021-00270-2.

UNICEF. 2020. *Draft report on policy guidance on ai for children*. https://www.unicef.org/globalinsight/media/1171/file/UNICEF-Global-Insight-policyguidance-AI-children-draft-1.0-2020.pdf.

Whittaker, Meredith, Kate Crawford, Roel Dobbe, Genevieve Fried, Elizabeth Kaziunas, Varoon Mathur, Sarah Mysers West, Rashida Richardson, Jason Schultz & Oscar Schwartz. 2018. *Ai now report 2018*. New York: AI Now Institute at New York University.

Zou, James & Londa Schiebinger. 2018. *Ai can be sexist and racist—it's time to make it fair*.

Chapter 3

Biases in AI in education

Leonie Grafweg, Lara-Sophie Witt & Paula Heupel

An algorithmic bias is the systematic and unfair discrimination of certain individuals or groups by computer systems. There are many different forms of biases like preexisting or technical biases. The affected groups are on the one hand commonly studied demographic categories like gender and nationality and on the other hand categories with limited research so far like disabilities or the parental educational background. Human biases are actually not that different from algorithmic biases. Both underlie similar procedures and both are not that easy to eliminate. The proposed solutions emphasize as a first step to uncover biases to understand and tackle the problem. Therefore we aim to spread public awareness about biases especially in education with this literature review.

Keywords: Bias, Education, Artificial Intelligence, AI

1 Introduction

When thinking about AI in education, a great amount of opportunities can come to one's mind. There can be computer systems supporting students, a lot of individual learning opportunities, and tools to help teachers decrease their workload. We realized that this is only a small fraction of all the possible applications when reading through the systematic research on AI in higher education by Zawacki-Richter et al. (2019). What attracted our attention was something that we felt was missing: ethical concerns were rarely considered, as Zawacki-Richter et al. (2019) stated in their paper. Another thing they mentioned was the lack of attention concerning risks and challenges, and the need for further research in the field of ethics of AI in education. We decided to put our focus on biases in AI in Education and will address the following questions with this paper:

- What is a bias?
- Why do biases arise?
- Who is affected by biases?
- Are there differences of biases in humans and biases in algorithms?
- Can biases in AIEd be prevented?

We conducted our search for relevant papers as described in Section 2. Some of the papers had their own definition of bias or referenced popular work on the topic bias in general. We gathered different definitions of bias in Section 3.1. Section 3.2 deals with different forms of biases and how they arise. We also took a closer look on the individuals and groups that are affected and disadvantaged by biases (Section 3.3). AI systems already start to take over tasks from human teachers. This could, for example, be witnessed in UK's A-Level scoring where an algorithm generated the students' grades. What happened was, that the algorithm assigned better grades to students from private or independent schools whereas students from rather disadvantaged backgrounds experienced a heavy downgrade (Smith 2020). It is therefore reasonable to compare biases in AI with biases in humans which can be problematic as well. We deal with that issue in Section 3.4. In order to not only talk about the problematic parts of the topic we address possible solutions in Section 3.5. We conclude the topic in Section 4.

2 Methodology

We conducted the search for applicable research papers by limiting our search to articles concerning biases in AI in education and biases in AI in general. We were hoping to be able to apply some of the broader bias challenges in AI to specific applications in the field of education. The articles used in our paper mainly originate from Scopus and Google Scholar, where we restricted the search to only open access documents with the keywords "Bias", "AI", and "Education". We did not limit our search to a certain time span since the topic of AI in education and possible biases is rather new in the research field and there is no final solution yet to avoid biases in AI. The publishing dates of the reviewed articles vary form 1996 to 2021. Across the databases we found a variety of documents including the defined keywords. Scopus left us with 24 open access documents including the keywords and Google Scholar found about one million results, of which we could not review all. Based on their title and abstract, we decided to select or not select articles for our review and ended up with 21 relevant articles, that we read and considered appropriate for the topic. A few other papers were brought to our attention by references of the papers we found when conducting our search.

Table 1: Criteria for our search

Inclusion criteria	Exclusion criteria
Keywords: Bias, AI, Education	No connection to AI or education
English or German language	Other languages
Open access	Limited access
Published 2021 and earlier	Published after 2021

The most challenging part of the search process was choosing relevant and irrelevant articles while not spending too much time reading through whole articles that would not find application in the end. Especially the Google Scholar database offered a large variety, that we could not go through exhaustively. We sorted the results by relevance and chose to occupy only those of notable pertinence. Furthermore, we were not able to include other than German or English articles due to our own language limitations.

3 Results

In this section we answer the questions posed in the introduction. We collect the different findings and definitions that could be found in the papers we found to be useful for our topic.

3.1 Definition of bias

When trying to analyze different biases in AI in education (AIEd) it is important to understand and clarify what is meant by the term "bias". It can, for example, be used in its pure statistical meaning or in situations where a bias does not cause any harm. On the other hand, the term bias can be used in situations where people with certain attributes suffer the consequences of unjustifiable decisions to their disadvantage in comparison to other people without these attributes. In order to not confuse different meanings and uses of the term bias we will present different definitions that were provided in the papers we deemed suitable for the topic.

One definition that is often referenced in literature on bias in AIEd is by Friedman & Nissenbaum (1996). In their paper, they define biased computer systems as "computer systems that systematically and unfairly discriminate against certain individuals or groups of individuals in favor of others" (Friedman & Nissenbaum 1996: p. 332). For this definition, it is important that a computer system does not randomly, but systematically, discriminate against individuals or groups. The computer system does not discriminate against a certain group on one day and against a different group on the next day. The same individuals or groups will always be affected by the computer system. The next important aspect of the definition is that the outcome of the discrimination is unfair. The affected individual or group has a disadvantage compared to non-affected individuals or groups. This disadvantage is based on "grounds that are unreasonable or inappropriate" (Friedman & Nissenbaum 1996: p. 332). Only when these two factors, systematic discrimination and an unfair outcome, occur together in a computer system, it is biased in the sense of Friedmann and Nissenbaum's definition.

Ferrer Aran et al. (2021) also differentiate between a broader definition of bias and the problematic kind of bias they refer to in their paper. Following their broad definition bias means "a deviation from the standard, sometimes necessary to identify the existence of some statistical patterns in the data or language used" (Ferrer Aran et al. 2021: p. 1). For the use of the term bias in their paper the authors specify their focus on "the problematic instances of bias that may lead to discrimination by AI-based automated-decision making systems" (Ferrer Aran et al. 2021: p. 1). Not defining the term bias further they still make sure to exclude biases that do not cause any harm from their definition.

When analyzing algorithmic-driven decision-making systems, Ferrero & Gewerc (2019) define identifying a bias as detecting "partial outputs that involve systematically unfair situations in which certain groups are benefited with privileges" (Ferrero & Gewerc 2019: p. 68). The aspects of systematic discrimination and an unfair outcome that Friedman & Nissenbaum (1996) explained in their paper can also be found in this definition. We found the definition of bias by Friedman & Nissenbaum (1996) to cover all the important aspects, so we decided to use the term in the sense they described.

Some authors prefer the term "unfair" over "biased" in order to use "bias" only in its statistical sense (Baker & Hawn 2021). We decided to use the term bias nevertheless. Substituting "biased" with "unfair" might imply that an unbiased algorithm is fair which is not necessarily the case (Silberg & Manyika 2019).

3.2 Different biases

We have now gathered different definitions of the term bias. Still, there are numerous different forms of biases. Similarly to the definition of the term, there is not one list of forms of biases but many definitions from different papers.

Friedman & Nissenbaum (1996) divided biases in three categories: preexisting bias, technical bias, and emergent bias. A preexisting bias is one that exists independently from the computer system. The bias already existed before the computer system was created. It does not matter whether this bias is an individual bias of a single person that has influence on the system or whether the origins of the bias lay in society as a whole. The bias can be included willingly with intent or unconsciously infiltrate the system even if one is determined to create a fair and unbiased algorithm.

A technical bias does not have its roots in an individual or societal bias but does arise from technical limitations that do not allow to obtain completely fair and equal results. Some examples for technical biases, given by Friedman & Nissenbaum (1996), were the following: alphabetically ranked results (giving disadvantage to results with a letter that is located towards the end of the alphabet), multiple screens of results one has to click through due to the size of a screen (favoring results that are displayed on the first screen where no further action is needed to view them) and, finally, imperfections in pseudorandom algorithms that are not truly random (e.g. favoring entries towards the end of a database when picking out one randomly).

An emergent bias arises mainly after a computer system is developed when it is already in use (Friedman & Nissenbaum 1996: p. 336). This can be the case when the environment, in which the system is used, differs from the intended environment or the actual users differ from the intended group of users. New knowledge or discoveries that are not considered by the system may cause such an effect. In the following we list some examples of emergent biases, given by the authors. Users with different knowledge or skills than those for which the system was designed may be disadvantaged, e.g., people with different language abilities using a language-based system. Computer systems used in a different culture with different values than the ones the system relies on may obtain very different results. Think of a gamified learning environment that works by competition being used with a cultural background where cooperation is highly promoted (Friedman & Nissenbaum 1996).

Ferrer Aran et al. (2021) also came up with three different categories of biases: bias in modeling, bias in training, and bias in usage. They describe biases in modeling as intended biases with the purpose of compensating for a bias in the data "through smoothing or regularisation parameters" (Ferrer Aran et al. 2021: p. 1).

A bias in training occurs when biases are represented in a data set that is used for training of an algorithm. If the training data is biased or not representative, the algorithm that is being trained absorbs this bias.

When algorithms are used for unintended purposes a bias in usage can arise. Misinterpretation of the output of an algorithm can also lead to this kind of bias.

Five different forms of biases were defined by Ferrero & Gewerc (2019). They classified different biases as follows: theoretical bias, methodological bias, bias by interpretation, bias by decontextualization, and bias by data training.

Theoretical biases can occur when complex concepts and phenomena are simplified too much in order to be able to represent them computationally. For example, social or cultural differences may be seen as a deficit in learning assessment because their results do not conform with the ones the AIEd system expects, although they are not false.

When data is not being handled with expertise or sufficient knowledge in statistics or programming it can happen that methodological biases, as defined by the authors, occur. Similar to the bias in usage described by Ferrer Aran et al. (2021), Ferrero & Gewerc (2019) define a bias by decontextualization which can arise when algorithms are used in contexts they were not designed for.

Generalizing an algorithm with the intent to use it in other contexts than originally planned can also lead to biases and mistakes.

The definition of bias by training data by Ferrero & Gewerc (2019) also coincides with the definition provided by Ferrer Aran et al. (2021). They compare a trained algorithm to a mirror that reflects the biases and false representations that may be contained in the training data. This bias often is a preexisting bias as described by Friedman & Nissenbaum (1996). The authors conclude by depicting that "algorithms are not biased by themselves; it is the [people creating and training the algorithms] who produce biases and decide, intentionally or not, to reproduce the injustices" (Ferrero & Gewerc 2019: p. 170).

3.3 Who is affected?

After we have clarified the terminology around biases and where they originate, we are moving on to finding out who is affected by biases in education. The affected groups are defined based on law-protected characteristics like in the Equality Act of 2010 in the UK. [1] There exist different law formulations around the world and there is no world standard for protected groups since minorities differ from nation to nation and are often originating from a national historical background. Baker & Hawn (2021) divided the target groups into legally defined classes, like ethnicity, nationality, and gender, and groups with limited research other than the major classes, like disabilities or parental educational background.

[1] https://www.legislation.gov.uk/ukpga/2010/15/contents

In the following, we are going to present some of the discovered biases for both the major and the rather unpopular demographic categories.

3.3.1 Commonly studied demographic categories

Ethnicity is a demographic category that is often studied, also in connection to bias in education. Anderson et al. (2019) conducted a study of a prediction model for six-year college graduation on different racial groups and five different algorithms while investigating the equity and optimality of the model. The results show a high false-positive rate for white students and a high false-negative rate for Latino and Hispanic students, meaning that white students were often inaccurately predicted to graduate and Latino and Hispanics were falsely assessed to not graduate, indicating gaps in the model's equity.

Besides ethnicity, nationality is a second commonly studied category. In an e-rater system study, automated essay scoring was compared to human essay rating by Bridgeman et al. (2009). The discrepancy between human and e-rater scores was notably large for Chinese and Korean students, indicating an upwards bias for those nationalities while predictions for other nations were accurate. These results were replicated a few years later and additionally, remarkably lower scores for Hindi, Arabic, and Spanish speakers were found for e-rater scoring compared to human scoring (Bridgeman et al. 2012).

The last major demographic category is gender. Anderson et al. (2019) did a study on a six-year college graduation prediction model, which showed higher false-negative rates for male students, indicating a downwards bias against them. Another study on an undergraduate grade prediction system showed inaccurate prediction for female students to perform better, indicating the opposing upwards bias for women (Yu et al. 2020). Baker & Hawn (2021) emphasize that studies were only conducted on the binary genders, male and female, and there are no studies on LGBTQ (Lesbian, Gay, Bisexual, Transgender, Queer) identity bias yet.

3.3.2 Limited research categories

There are a lot of limited research categories that would require deeper investigation and there are even categories that still cannot be classified at all.

One of these categories is for example the native language or dialect. Naismith et al. (2018) did a measurement study of lexical sophistication and found that measures were more accurate when frequency lists were based on the second-language-learners word pool than when frequency lists were based on native speakers. In terms of dialects, there are no significant studies on biases in AIEd.

A second category that is in need of deeper research is disability because a lot of work on bias is focused mainly on the detection of disabilities in general, not on biases against disabled people. AI systems are limited when it comes to recognizing body language and gestures of disabled people or in automatic speech recognition for atypical speech (Guo et al. 2020).

Yu et al. (2020) did a study on college success prediction and investigated the impact of the parental educational background on students, which forms another limited research category. First-generation college students were often underestimated and inaccurately predicted to perform worse.

There are further categories to list here (e.g. urbanity, socioeconomic status, military-connected status, age or religion) but this would go beyond the scope of this review.

3.4 Comparison: human biases vs algorithmic biases

Algorithms already start to take over tasks from humans, as stated in Section 1. Since these algorithms also were created by humans, it is reasonable to compare human biases and algorithmic biases.

We especially focus on the similarities and differences of human and algorithmic biases. The first similarity between the two is that human biases and algorithmic biases both mainly are about the same topics. Baker & Hawn (2021) named the major topics from algorithmic biases: ethnicity, nationality and gender. These are the main biases in humans as well. Of course, the topics of the biases are not exactly the same, but the parent topics are very similar. Another similarity of human biases and algorithmic biases is their origin. Mehrabi et al. (2021) claimed that algorithmic biases can develop, if users with preexisting biases make use of the computer system, their interactions are converted to data and the data then affects the algorithm. For humans, it is similar. Most biases from humans exist because someone told them about an experience or they made one by themselves, which is then converted into "data" and is influencing the humans in their later actions. So in both cases something happens, the information of this event is saved and is then responsible for future beliefs. For example, if your sibling had the same teacher in school and received bad grades, the teacher maybe is biased against your family and will give you bad grades as well. The same procedure applies to algorithmic biases, the computer got similar input data because you and your sibling have a similar social background, it might rate your tasks worse, than without this information.

This leads to the first difference in comparison of human biases and algorithmic biases. If algorithmic biases were recognized, one can remove them by modeling the data sets (Fazelpour & Danks 2021). For that, one should avoid sensitive data about group memberships. For example, if bias about race in school is not wanted, one should remove racial identifiers. The problem is with this procedure one cannot guarantee for a reliable program anymore because for some actions maybe information is missing then. Human biases cannot be eliminated so easily. If a human made an experience, which lead to a bias, the experience cannot just be removed. To eliminate human biases one needs to be convinced that the bias is not appropriate, which is way harder. Another difference exists with regard to technical biases which arise due to technical limitations. Humans do not have these limitations in comparison to computer systems and therefore do not have technical biases either.

A difference is also the diversity in the used data. Humans make experiences or get information from other people or literature for their "data sets". Computer systems on the other hand receive information for their "data sets" from everywhere. Of course, they get information from the developers, the internet, literature, and other sources. Humans have access to these sources as well, but in a much smaller range than computer systems. For example, a computer receives information from the developers of course and every other user. A human is in contact with way fewer people, namely just people in their vicinity. For that reason is the data from computer systems much more diverse.

3.5 Possible solutions

We outlined how bias is defined and who is affected by it in an educational context. Now we want to focus on possible solutions to reduce bias in AIEd and we focus on the obstacles and their resolutions presented by Baker & Hawn (2021). The basic strategy, one should try to follow, is moving from unknown biases to uncovering them and from the known biases to tackling the problem. This allows to address the causes effectively and to not spend time speculating about the symptoms. There will not be an optimal state where no algorithmic bias can be found, but we can strive for increasing fairness as good as possible. The challenges that need to be addressed start with revealing the biases and understanding their causes. Secondly, it is crucial to collect more representative data for the training of models, but the data collection process often raises privacy concerns. The goal is "to balance the risk of privacy violations with the risk of algorithmic bias" (Baker & Hawn 2021). There may also be economic obstacles that prevent researchers from publicly presenting biases in their algorithms. They would risk public critique or lawsuits and therefore avoid attracting attention by publishing bias evidence to retain a certain reputation. Baker & Hawn (2021) claim that society needs to demand evidence for bias-free education, from humans as well as in algorithms. The recommendations for the improvement of biased algorithms start with improving the data collection process by making appropriate data available and requiring diverse demographic data. Baker & Hawn (2021) also point out that the current research has a mainly American focus, which emphasizes the necessity for research in larger international contexts beyond the US. In the collection process, one should also be cautious about collecting already biased training data and avoid its use. Furthermore, an improvement of tools and resources is achieved by making standard tests and guidelines for biases available. Finally, we could increase fairness by broadening the community and including everyone who is affected by biases in some way. We should make an effort to explain AIEd and spread awareness in order to include not only academics since AI is probably going to affect most of us.

4 Conclusion

To conclude, biases of computer systems arise because they have technical limitations, the computer system was used by someone with biases, or the developers already had biases when they developed the computer system.

Most of the times the people who are affected by biases in education belong to one of the legally protected classes, like ethnicity, nationality, or gender, and attributes with limited research like disabilities or parental educational background. These biases can be quite similar to known human biases. To abolish these biases, we should uncover unknown biases and tackle the problems of known biases. In the end, it is quite unlikely that biases will disappear completely. For that reason it is important to take action and inform people about biases. Everybody who works with AI should know about biases and be able to handle them. We want to emphasize that there are inconsistent findings across different studies and there is a need for a larger number of studies on a larger range of contexts. Nevertheless, it should be the ultimate goal to improve computer systems in a way that as many biases as possible dissolve.

References

Anderson, Henry, Afshan Boodhwani & Ryan S. Baker. 2019. Assessing the fairness of graduation predictions. In Michel C. Desmarais, Collin F. Lynch, Agathe Merceron & Roger Nkambou (eds.), *Proceedings of the 12th international* *conference on educational data mining, EDM 2019, Montréal, Canada, july 2-5, 2019*, 488–491. Worcestor: International Educational Data Mining Society.

Baker, Ryan S. & Aaron Hawn. 2021. Algorithmic bias in education. *International Journal of Artificial Intelligence in Education* 2021. Published 18 November 2021. DOI: 10.1007/s40593-021-00285-9.

Bridgeman, Brent, Catherine Trapani & Yigal Attali. 2009. Considering fairness and validity in evaluating automated scoring [paper presentation]. https://www.researchgate.net/publication/242203403_Considering_Fairness_and_Validity_in_Evaluating_Automated_Scoring.

Bridgeman, Brent, Catherine Trapani & Yigal Attali. 2012. Comparison of human and machine scoring of essays: Differences by gender, ethnicity, and country. *Applied Measurement in Education* 25(1). 27–40. DOI: 10.1080/08957347.2012.635502.

Fazelpour, Sina & David Danks. 2021. Algorithmic bias: Senses, sources, solutions. *Philosophy Compass* 16(8). DOI: 10.1111/phc3.12760.

Ferrer Aran, Xavier, Tom van Nuenen, Jose M. Such, Mark Coté & Natalia Criado. 2021. Bias and discrimination in AI: A cross-disciplinary perspective. *IEEE Technology and Society Magazine* 40(2). 72–80. DOI: 10.1109/MTS.2021.3056293.

Ferrero, Federico & Adriana Gewerc. 2019. Algorithmic driven decision-making systems in education: Analyzing bias from the sociocultural perspective. In Lisa O'Conner (ed.), *Proceedings - 14th Latin American conference on learning technologies, LACLO 2019*, 166–173. San Jose del Cabo: Institute of Electrical & Electronics Engineers Inc. DOI: 10.1109/LACLO49268.2019.00038.

Friedman, Batya & Helen Nissenbaum. 1996. Bias in computer systems. *ACM Trans. Inf. Syst.* 14(3). 330–347. DOI: 10.1145/230538.230561.

Guo, Anhong, Ece Kamar, Jennifer Wortman Vaughan, Anna M. Wallach & Meredith Ringel Morris. 2020. *Toward fairness in AI for people with disabilities: A research roadmap*. New York, NY, USA. DOI: 10.1145/3386296.3386298.

Mehrabi, Ninareh, Fred Morstatter, Nripsuta Saxena, Kristina Lerman & Aram Galstyan. 2021. A survey on bias and fairness in machine learning. *ACM Comput. Surv.* 54(6). 1–35. DOI: 10.1145/3457607.

Naismith, Ben, Na-Rae Han, Alan Juffs, Brianna Hill & Daniel Zheng. 2018. Accurate measurement of lexical sophistication in ESL with reference to learner data. In Kristy Elizabeth Boyer & Michael Yudelson (eds.), *Proceedings of the 11th international conference on educational data mining, EDM 2018, Buffalo, NY, USA, july 15-18, 2018*, 259–265. Worcestor: International Educational Data Mining Society (IEDMS). https://educationaldatamining.org/EDM2018/.

Silberg, Jake & James Manyika. 2019. *Notes from the AI frontier: Tackling bias in AI (and in humans)*. Chicago. https://www.mckinsey.com/featured-insights/artificial-intelligence/tackling-bias-in-artificial-intelligence-and-in-humans (2 December, 2021).

Smith, Helen. 2020. Algorithmic bias: Should students pay the price? *AI & SOCIETY* 35(4). 1077–1078. DOI: 10.1007/s00146-020-01054-3.

Yu, Renzhe, Qiujie Li, Christian Fischer, Shayan Doroudi & Di Xu. 2020. Towards accurate and fair prediction of college success: Evaluating different sources of student data. In Anna N. Rafferty, Jacob Whitehill, Violetta Cavalli-Sforza & Cristobal Romero (eds.), *Proceedings of the 13th international conference on educational data mining (EDM 2020)*, 292–301. Virtual event: Proceedings of The 13th International Conference on Educational Data Mining (EDM 2020). https://educationaldatamining.org/edm2020/proceedings/.

Zawacki-Richter, Olaf, Victoria I. Marín, Melissa Bond & Franziska Gouverneur. 2019. Systematic review of research on artificial intelligence applications in higher education–where are the educators? *International Journal of Educational Technology in Higher Education* 16. 39. DOI: 10.1186/s41239-019-0171-0.

Chapter 4

Consequences of AI for teachers in K-12 education

Kyra Breidbach & Antonella Rönck

This systematic literature review analyzed 15 papers in respect to challenges and opportunities that artificial intelligence (AI) and the use of AI brings upon teachers in K-12 (kindergarten through high school) education. While investigating the papers, we mainly focused on answering the questions what challenges teachers face to integrate AI into the teaching practice and to what extent AI and its applications can be helpful to them. We found that educators struggle with acquiring more complex technological skills, the change of their role as a teacher, and facing ethical challenges. However, many applications enable and support personalized teaching methods. Another opportunity of human-machine collaboration is saving valuable teaching time and also improving the teaching quality.

Keywords: teacher assistance, K-12, challenges, opportunities, personalized learning, human-machine collaboration

1 Introduction

The development of Artificial Intelligence in education (AIEd) provokes research on how the employment of AI-driven technologies effects educators. In this literature review we focused on the consequences teacher in K-12 education face in the light of use of AIEd.

AIEd provides support for teachers in numerous ways. While this can be an opportunity to solve various challenges, training teachers in learning how to handle AIEd applications in general and with respect to their ethical use can be very time consuming. Though, repetitive and tedious tasks previously executed by the educators can consequently be automated and save teachers this valuable time.

In the context of education, AI offers multiple solutions to support teachers and enables individualized learning for the students (Chounta et al. 2021). Simultaneously, the opportunity to redefine the role of the teacher in a classroom setting from the sole source of knowledge to a collaboration between humans and machines will be a great challenge (Liu & Wang 2020).

This paper will describe those opportunities and challenges that AI poses to teachers in K-12 education, how AI impacts their role, and in which way this role needs to be redefined. While answering those questions, the paper will solely focus on the teachers' perspective and not take into account the ways the society and educational system as a whole could support them. It will also not specify the consequences AI brings onto the school system as a whole or the students, nor will it cover the issue of AI as a replacement for teachers. The focus will be on how AI can support teachers in K-12 school systems in their work and how teachers will need to learn to cooperate with AI-based systems.

2 Methodology

A systematic review aims to answer specific questions based on scientific reports and it needs to be conducted in an explicit, systematic and replicable manner (Zawacki-Richter et al. 2019). To fulfill these requirements a systematic review openly displays inclusion and exclusion criteria that are crucial for data collection. This systematic literature review encompasses fifteen papers on the consequences of artificial intelligence for teachers in K-12 education.

2.1 Search strategy

The search consisted of the strings "Artificial" AND "intelligence" AND "in" AND "education" AND "teacher" over the databases Scopus, Google Scholar and PubMed. We also used additional strings and further specification of our search terms for each database can be found in Table 1.

The search included articles written in English and was conducted in December 2021. Since the field of AI is developing rapidly we decided to exclusively take a look at relatively new papers to ensure that our findings are of current importance. Thus our literature research was limited to publications of the last five years and those that are dated to be published in 2022. Across all databases the results were sorted by relevance. While databases such as Scopus and PubMed found a reasonable amount of papers for the given search strings, Google Scholar presented significantly more results. Therefore only the top thirty documents for each individual search string were included in the initial literature research.

Once every search string had been entered across the three databases 331 documents were found. A total of 93 duplicates could be removed leaving 238 documents for screening the papers' abstracts and titles. If a document fulfilled at least one of the exclusion criteria and none of the inclusion criteria, which can both be found in Table 2, the document was excluded. However, if a title or an abstract did also satisfy one or more of the inclusion criteria and there was still a possibility that the paper was not indefinitely fully about something listed as an exclusion criterion, we decided to keep it for a thorough reading. In this step 212 documents were eliminated. Afterwards, the remaining 26 documents were checked for retrievability. Another six documents were excluded in this step. After fully reading twenty papers, another five documents were excluded as they were deemed irrelevant to the topic of this review according to our exclusion criteria (Table 2). Four of the final fifteen papers that were analyzed were literature reviews.

The process of our paper collection can be found in Figure 1.

Table 1: Search strings used for literature research

Database	Specified search string
Scopus	”K-12” or ”challenge” or ”opportunity”
PubMed	”K-12” or ”challenge” or ”opportunity”
Google Scholar	”K-12” and (”challenge” or ”opportunity”)

Table 2: Criteria for excluding papers from the review

Inclusion criteria	Exclusion criteria
Published 2017-2022	Published before 2017
English language	Not English language
K-12 or general education	Higher education or other specific education (e.g. medical)
Consequences for teachers	Consequences for students or educational institutions
Use of AI in education	AI as a subject
	AI as replacement of teachers not for support
	No AI in education
	Applications of AI for teachers

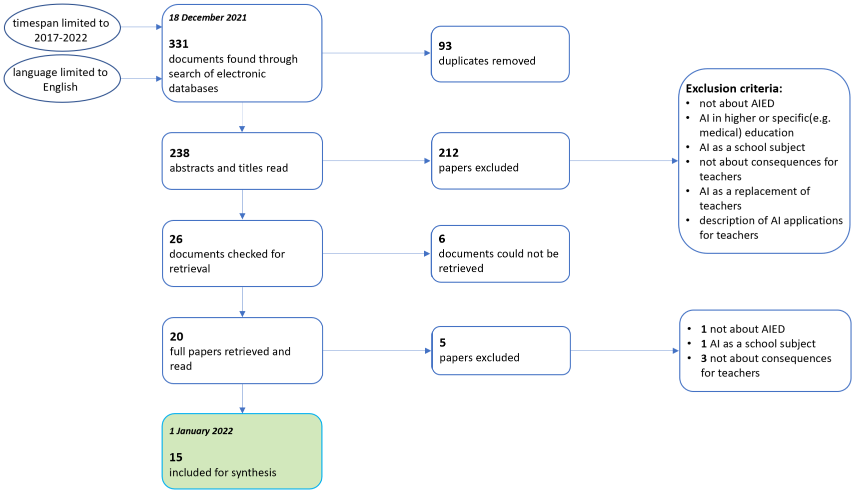


Figure 1: Process of paper collection

2.2 Limitations

A limitation of every literature review is the applied search strategy. While maximizing the results achievable with the given resources, the search was limited by the three databases used. When using the database Google Scholar, we decided to include only the top thirty articles into our initial search. This measure was taken due to time restraints but also limits the choice of possible papers for our review. Furthermore, research not conducted in English could not be included. Neither could articles that were not openly accessible nor retrievable via the institution Osnabrück University.

Since some interesting papers were reviews, we decided to include every kind of document that we found via the previously described search strategy into our literature review. Therefore, since we did not only look into original studies, but also into other researchers' reviews, our findings might be slightly biased to the reviews influences. Lastly, while we excluded papers that explicitly stated to not cover K-12 education, papers were included even if they did not explicitly state that they were analyzing K-12 education.

3 Results

In the following section the papers are further analyzed regarding the countries of publication and the author affiliations as well as the findings about opportunities and challenges AI yields for teachers. While some authors greatly glorified the usage of AI and others concentrated on possible ethical issues, the majority of results remained balanced.

3.1 Countries

For the determination of the papers' geographical origination we looked into the country of the first authors workplace or the place of establishment of the company which was done for the paper by Bryant et al. (2020). Most papers came from US authors (n = 5) and the second most originated from Chinese (n = 4) researchers. Four of the papers were published by European associations, two of these papers came from Spain, another from Estonia and the last from Sweden. Canada and the United Arab Emirates each contributed one paper as well.

3.2 Author affiliations

Similar to the previous section, only the first author's affiliation was taken into consideration. The majority of the analyzed papers was written by authors working in university departments of education (n = 7). Two papers were contributed from technological facilities, one being from a? Department of Learning Technologies and another from a? Division of Digital Learning from the Institute of Technology. One paper was published in a university's Computer and Information Science Department and another at a Human-Computer Interaction Institute. We also included two papers by the university facilities School of Journalism and Communication as well as School of Humanities. Moreover, one paper was contributed by an Economics and Business Foreign Language School. Only one of the sources included was published outside of a university context but within the US-American based business McKinsey.

3.3 Opportunities

According to Liu & Wang (2020) AI is the driving force of the classroom reform.

The most prominent advantage of AI for teachers is that instead of replacing them it could lead to a human-machine collaboration (Xia et al. 2022, Murphy 2019, Liu & Wang 2020, Buliva n.d.) that will save the teachers time through completing repetitive and time intensive tasks and thereby improve the teaching quality (Xia et al. 2022, Murphy 2019, Liu & Wang 2020, Vázquez-Cano 2021, Buliva n.d., Bryant et al. 2020, Pedró et al. 2019, Wang 2021). For example, AI can provide immediate results with a high accuracy both as feedback to the student and information about the students progress for the teacher (Xia et al. 2022, Murphy 2019, Liu & Wang 2020, Buliva n.d., al-Zyoud 2020, Akgun & Greenhow 2021, Chounta et al. 2021).

Specific-purpose AI applications, such as language translation applications, could be used to improve professional abilities (Chounta et al. 2021) such as being able to use material offered in different languages and exchanging knowledge with teachers form other countries.

AI facilitates the design of learning activities as it can be used to organize and plan lessons in regard to schedule and contents (Chounta et al. 2021). It can identify struggling students and intervene to prevent drop-out (Murphy 2019, Liu & Wang 2020, Vázquez-Cano 2021, Buliva n.d., Holstein et al. 2019), and by shifting the teachers' role away from the knowledge holder it can free time for the teacher to support the students (Akgun & Greenhow 2021).

The evaluation provided by AI allows teachers to focus less on the results and more on the process such as acting as the students' learning partners and motivators (Xia et al. 2022, Holstein et al. 2019). While the educators remain the decision makers and leaders (Liu & Wang 2020), AI co-orchestration systems help achieve complex and powerful learning scenarios (Holstein et al. 2019) that prepare students for the digital age (al-Zyoud 2020), since AI supports students learning experience and teachers' practice (Akgun & Greenhow 2021). Moreover, real-time assessment applications could show what might help individually as well as enable teachers to get an insight on students' cognitive states (Chounta et al. 2021) and to better support students when needed (Wang 2021).

Furthermore, AI could have the potential to lead to worldwide collaboration to make information accessible such as educational plans or educational outcomes (Pedró et al. 2019) and thereby improve teaching styles independent of one's location.

Another outstanding opportunity AI offers is the individualization of education. It gives the possibility to teach in accordance with each students' knowledge state through self-pacing (Xia et al. 2022, Vázquez-Cano 2021, al-Zyoud 2020, Hrastinski et al. 2019, Pedró et al. 2019, Wang 2021, Chounta et al. 2021) and thereby gives differentiated and more effective instructions in mixed-ability classrooms (Murphy 2019, Liu & Wang 2020, Akgun & Greenhow 2021).

One example for this is analyzing students' fitness level and adapting physical education accordingly (Wang 2021). This individualized learning results in an education similar to one-on-one teaching with improved learning outcomes and students' achievements (Murphy 2019, Chounta et al. 2021). Furthermore, it leads to students requiring less assistance and henceforth saves even more of the teachers' time (Buliva n.d.).

The saved time will prevent teachers from leaving their work due to workload (Bryant et al. 2020). Additionally, it enables them to focus on their responsibility to teach the students those things that AI cannot convey like social and emotional skills such as empathy and creativity (Murphy 2019, Holstein et al. 2019).

It changes their role from being responsible for "imparting knowledge to guiding the healthy growth of students" (Wang 2021: 952). With low-level abilities like memory and recitation being replaced by the digitalization and therefore less necessary to teach, education in K-12 will focus on the development of personality and innovation ability that enables students to apply and analyze knowledge (Liu & Wang 2020, Wang 2021). The increased usage of AI in- and outside of education forces teachers to emphasize authentic and project-based learning activities which changes their task "from preparation of students to entering a workforce to readying them to become adaptive experts and on-the-job learners" (Buliva n.d.: 3). Teachers become mostly responsible for guidance (Wang 2021). They remain substantial for the internal regulations in regard to social communication and creative aesthetics (Liu & Wang 2020) as well as higher order critical thinking, common sense (Murphy 2019), and as their students' educator and leader (Liu & Wang 2020).

Thereby, the integration of AI into the educational environment brings various opportunities into the teachers' didactic practice. The applications that AI enables will benefit both the educators and the students.

3.4 Challenges

A difficulty that will arise for teachers is the integration of AI into the current curriculum and the thereby necessary task change. Since students progress at different rates, AI will be hard to incorporate into lessons (Murphy 2019). This is especially true since willingness to implement AI technology is not an easy task for teachers (Bryant et al. 2020). Teachers are required to be more innovative in their usage of AI to change teaching ideas and methods (Wang 2021) and simultaneously rethink educational goals (Liu & Wang 2020). Accordingly, graduating students no longer need to possess a maximum amount of knowledge but instead they need to know how to navigate resources and apply their search-results. Another related challenge is the need to choose among a myriad of competing AI solutions when designing the lesson plan (Bryant et al. 2020). Over-automation of education might take over classroom roles and therefore threaten flexibility, while under-automation will burden teachers with avoidable tasks and limit their degree of personalization (Holstein et al. 2019). Teachers have to realize cross-regional and cross-domain teaching (Wang 2021) to be able to exchange knowledge and experience about teaching with AI internationally and across different subjects. Therefore, they need to enter a new dialog with educators, content designers, and cross-disciplinary specialists (Pedró et al. 2019).

With the integration of AI into the educational system, wider and more complex technical skills such as digital competence are required from teachers (al-Zyoud 2020, Hrastinski et al. 2019, Pedró et al. 2019, Wei et al. 2020). They will need continuous training (Liu & Wang 2020, al-Zyoud 2020, Chounta et al. 2021, Wei et al. 2020, Adams et al. 2021) to learn how to use AI teaching tools (Wang 2021) and interpret it (Chounta et al. 2021). While they develop their own literacy, they also have to motivate other teachers to do the same (Wei et al. 2020). The diversity that AI enables depends upon good resource design, integration capabilities, flexibility, and careful monitoring of students (Murphy 2019, Wei et al. 2020).

In order to meet these requisites teachers' role changes away from the monopoly of knowledge (Liu & Wang 2020). Education shifts from teacher-centered to student-centered and personalized (Wei et al. 2020). While this entails diverse opportunities once achieved, the adjustment will be a challenge. The teacher has to be "the engineer of the student soul, [...] the AI applicator" (Wei et al. 2020: 718), the organiser of learning, and the cultivator of emotional values (al-Zyoud 2020, Chounta et al. 2021). Therefore, teachers need to find new balance in their sense of identity (Hrastinski et al. 2019, Wei et al. 2020).

Another common worry is that the relationship to students might be affected as AI could potentially "hinder social aspects of learning" (Chounta et al. 2021: 20) such as communication. According to Chounta et al. (2021) 'real' (that is, human to human) communication cannot be substituted by human-machine communication.

This leads to the final challenge of AI for teachers; ethics. When applying AI in education settings teachers have to ensure ethical use, be aware of learned bias and prevent its influence as much as possible. After being trained AI will reconstruct the unfair and inconsistent results given in its training data and teachers have to undermine those stereotypes and potential racial discrimination possibly even more than their own (Murphy 2019, Akgun & Greenhow 2021, Chounta et al. 2021, Adams et al. 2021). Another challenge that arose after Holstein et al. (2019) interviewed selected teachers is that they do not have any desire to mediate between students and AI as they view this as a waste of their time. They would rather trust the system to fairly and in an unbiased way decide which student has passed their respective level. This leads to issues with accountability and responsibility when AI makes mistakes such as discrimination since biases can oftentimes be unperceivable for the teacher. Moreover, the lack of autonomy raises the question of who is right when AI and teacher disagree (Hrastinski et al. 2019, Chounta et al. 2021, Adams et al. 2021).

Furthermore, integrating AI into the school system and thereby forcing students to use it leads to problems with the protection of their privacy and surveillance (Akgun & Greenhow 2021, Adams et al. 2021). Additionally, there might be an exacerbation of the digital divide not only for individual students who might not have access to computers but also in comparison to developing countries who have no possibility to use the AI systems (Pedró et al. 2019, Adams et al. 2021).

Business interest might set the agenda for the role and agency of teachers, while this agency could further diminish as AI usage might reduce introspective and independent thought (Akgun & Greenhow 2021).

Lastly, while integrating AI into lessons pedagogical appropriateness and thereby child centeredness need to remain in focus, as well as teachers well-being. There needs to be more awareness about increased workload, changing working conditions, additional time for preparation, shifting relationships with students, and worries about technological unemployment (Adams et al. 2021).

Concluding, with the integration of AI into the teaching practice teachers need to face various challenges concerning their skill requirements, the change of their role, the incorporation of AI into their curriculum, the change of human interaction, and ethical concerns.

4 Conclusion

In summary, AI entails various opportunities and challenges for teachers. With the introduction of AI into the educational system both teachers and students will profit from human-machine collaboration. The time teachers save with AI taking over repetitive tasks can be used to further guide students in building their social skills and personality. Additionally, AI improves the education through designing learning activities and preventing drop-out. This education can be individualized to teach in accordance with each students ability.

However, AI also places higher requirements on teachers who need to learn how to teach with and through it. They have to change their role as a teacher and adaptively react to the new teaching environment. At the same time, teachers need to ensure that ethical standards as well as their students' alike their own well-being will be protected. Integrating AI into the existing system will be a great challenge for teachers to face and it is one that requires a lot of additional training and willingness.

Concluding, teachers and the K-12 educational system in general will have to adapt to a variety of new possibilities. While most of them remain promising, the challenges that are destined to come should not be ignored. Regardless, the future development of AI will bring many advantages for teachers and students and should be greatly supported. Accordingly, further research in this field should be conducted to dynamically observe the effects that AI has on educators in order to potentially develop new AIEd technologies and make education as feasible and effective as possible.

References

Adams, Cathy, Patti Pente, Gillian Lemermeyer & Geoffrey Rockwell. 2021. *Artificial intelligence ethics guidelines for k-12 education: a review of the global landscape* (Artificial Intelligence in Education). Utrecht: Proceedings. 24–28. DOI: 10.1007/978-3-030-78270-2\{_}4.

Akgun, Selin & Christine Greenhow. 2021. Artificial intelligence in education: addressing ethical challenges in k-12 settings. *AI Ethics*. 1–10. DOI: 10.1007/s43681-021-00096-7.

Bryant, Jake, Christine Heitz, Saurabh Sanghvi & Dilip Wagle. 2020. How artificial intelligence will impact k-12 teachers. *McKinsey |& Company*. 1–7.

Buliva, Newton. N.d. Evaluating the role of artificial intelligence in k-12 education. a review of literature.

Chounta, Irene Angelica, Emanuele Bardone, Aet Raudsep & Margus Pedaste. 2021. Exploring teachers' perceptions of artificial intelligence as a tool to support their practice in estonian k-12 education. *International Journal of Artificial Intelligence in Education*. 1–31. DOI: 10.1007/s40593-021-00243-5.

Holstein, Kenneth, Bruce McLaren & Vincent Aleven. 2019. *Designing for complementarity: teacher and student needs for orchestration support in ai-enhanced classrooms* (Artificial Intelligence in Education). Cham: Springer. 157–171. DOI: 10.1007/978-3-030-23204-7\{_}14.

Hrastinski, Stefan, Anders D. Olofsson, Charlotte Arkenback, Sara Ekström, Elin Ericsson, Göran Fransson, Jimmy Jaldemark, Thomas Ryberg, Lena-Maria Öberg, Ana Fuentes, Ulrika Gustafsson, Niklas Humble, Peter Mozelius, Marcus Sundgren & Marie Utterberg. 2019. *Critical imaginaries and reflections on artificial intelligence and robots in postdigital k-12 education*. Vol. 1 (Postdigital Science and Education). Cham: Springer. 427–445. DOI: 10.1007/s42438-019-00046-x.

Liu, Jing & Shuli Wang. 2020. *The change of teachers' role in teaching under the environment of "artificial intelligence +"* (Proceedings - 2020 International Conference on Artificial Intelligence and Education, ICAIE 2020). Tianjin: IEEE Xplore. 98–102. DOI: 10.1109/ICAIE50891.2020.00030.

Murphy, Robert. 2019. Artificial intelligence applications to support k-12 teachers and teaching a review of promising applications, opportunities, and challenges. DOI: 10.7249/PE315.

Pedró, Francesc, Miguel Subosa, Axel Rivas & Paula Valverde. 2019. *Artificial intelligence in education: challenges and opportunities for sustainable development*. Vol. 7 (Working Papers on Education Policy). Paris: UNESCO.

Vázquez-Cano, Esteban. 2021. Artificial intelligence and education: a pedagogical challenge for the 21st century. *Educational Process International Journal* 10(3). 7–12. DOI: 10.22521/edupij.2021.103.1.

Wang, Chuanju. 2021. *The value orientation of teachers' role in artificial intelligence teaching environment based on information technology* (4th International Conference on Information Systems and Computer Aided Education). New York: Association for Computing Machinery. 951–954. DOI: 10.1145/3482632.3483059.

Wei, Qingmin, Mingyong Li, Kaiyue Xiang & Xue Qiu. 2020. Analysis and strategies of the professional development of information technology teachers. *15th International Conference on Computer Science and Education*. 716–721. DOI: 10.1109/ICCSE49874.2020.9201652.

Xia, Jinling, Huafeng Liu & Wenbin Liu. 2022. *Ai-based iwrite assisted english writing teaching*. Vol. 98 (Lecture Notes on Data Engineering and Communications Technologies). Cham: Springer. 158–165. DOI: 10.1007/978-3-030-89511-2\{_}19.

Zawacki-Richter, Olaf, Victoria I. Marín, Melissa Bond & Franziska Gouverneur. 2019. Systematic review of research on artificial intelligence applications in higher education–where are the educators? *International Journal of Educational Technology in Higher Education* 16. 39. DOI: 10.1186/s41239-019-0171-0.

al-Zyoud, Hashem Mahmoud Muslim. 2020. The role of artificial intelligence in teacher professional development. *Universal Journal of Educational Research* 8. 6263–6272. DOI: 10.13189/ujer.2020.082265.

Chapter 5

AI teacher assistants

Fabienne Kock, Anna Sommer

AI assistance can be a huge benefit for teachers in and outside the classroom, though the approaches diverge. In this paper we briefly take a look at the current research on AI assistance for teachers in education by conducting a small literature review on papers published after 2010. Afterwards we expand on three types of applications: Lumilo, Chatbots for language learning and Learning Analytics Dashboards. We found that AI teacher assistance is a growing field in the context of AI in Education. In general the assistance either gives the teacher (immediate) feedback on students and their performance, to identify situations where further guidance is needed or is a tool to predict a students performance based on existing data.

Keywords: teacher assistant, artificial intelligence, Lumilo, chatbots, learning analytics dashboard

1 Introduction

The data gathered on learning behavior cannot only profit the student themselves but can also enable the teacher to adapt to the students' needs. But teachers have different needs for AI in education compared to students. Besides tracking the students' progress, they generally use further verbal and non-verbal cues to ascertain the way in which they can optimise the individuals' progress. Taking these requirements into account different tools have been developed to assist teachers both in and outside the classroom.

Taking a closer look at a small selection of literature, we aim to give a general idea on the current state of research on AI tools as teacher assistance. Afterwards, we introduce three different kinds of application that stood out the most to us. Lumilo, a pair of smart glasses, that uses augmented reality to provide the teacher with real time analytics on their students, Replika, an application supplementing language learning via a chatbot and the general tool of Learning Analytics Dashboards providing a visual representation of learning analytics.

2 Method

To get a general overview of the field of AI teacher assistance we conducted a small literature review.

2.1 Search strategy

To get a general idea of applications implementing AI assistance for teachers in and outside the classroom, the initial search included articles published since 2010 that can be found using a set of predefined search strings (Table 1) in the two databases Google Scholar and Scopus. Since search using Google Scholar led to an excessive amount of papers, we limited us to the 20 first results found for the search strings. This initial corpus was screened for peer-reviewed, primary research articles published in English on applications assisting teachers in and outside the classroom using AI. Duplicates were removed. The initial search was conducted in December 2021 with a result of 86 articles. Based on the findings on AI assistance for teachers, three applications (Lumilo, Replika, intelligent analytics dashboard) were chosen to expand upon.

Table 1: Set of initial search strings

search string	total x of papers	x of usable papers
learning analytics AND teacher AND AI	17 scopus, 20 gScholar	7 scopus, 4 gScholar
ai AND teacher AND assistant	33 scopus, 4 gScholar	2 scopus, 0 gScholar
ai AND supported AND teacher	10 scopus, 2 gscholar	2 scopus, 0 gscholar

2.2 Limitations

The initial literature search is rather limited in search terms an scope. Though we tried to diversify using several search strings, the strings themselves were restricted to a small sub-field and do not cover the full breadth of AI assistance for teachers. Since we are limited in time and manpower to conduct an in depth literature review, we tried to compensate by giving a small overview on the information we found, before expanding on three specific applications.

3 General overview

AI assistance for teachers is a currently rapidly growing field. Most papers we found were from 2019 to 2021. Studies tested applications in all levels of schooling beginning with primary school (Michos et al. 2020), though most studies happen at university level. In our limited scope we have nearly as many tools assisting in online teaching as tools aiding in the classroom. Most classroom tools make use of a blended learning approach, merging online and traditional learning.

The type of assistance is actually quite diverse. There are external tools like the Lumilo smart glasses (Holstein et al. 2019b) or a robot (Hsieh et al. 2020) to assist teachers in class, as well as integrated tools tracking and predicting the learning progress of students enabling teachers to adapt their lecture content and offered online material (Rincón-Flores et al. 2020, Majumdar et al. 2021) to the students' needs. One commonality of many papers is an anayltics dashboard to visualize the data for the teacher. To accommodate the teachers' requirements for the visualization of the analytics, already a lot of newer research, similar to Holstein et al. (2019a) for Lumilo, already include teachers as stakeholders in the design process (Michos et al. 2020, Franzoni et al. 2020). Overall it can be said, that while there are already several attempts to successfully design AI assistance for teachers there is still much room for improvement. This will be discussed in detail in the following section of application examples.

4 Application examples

We choose three applications to showcase the developing field of AI as teacher assistants. We start with the analytic glasses Lumilo, that help in classroom orchestration, followed by the application of chatbots with Replika to aid teachers in language learning scenarios. Finally, we showcase the general idea of learning analytics dashboards as a tool to visualize the data gathered in a classroom.

4.1 Lumilo

First let us look at an analytics tool designed to assist teachers in the classroom: the analytics glasses Lumilo. Lumilo was designed by Kenneth Holstein and his colleagues to provide help to teachers to identify problematic situations in the classroom (Holstein et al. 2019a). In their essence the glasses were designed, so that teachers could get supported by educational technologies, such as intelligent tutoring systems (ITSs) in their effectiveness. They are supposed to „augment teachers state awareness during ongoing learning activities“ (Holstein et al. 2019b)[p. 1] by allowing the teachers to access real-time analytics on „student knowledge, progress, metacognition, and behavior within educational software“ (Holstein et al. 2019b)[p. 1] and to improve not just students' performance, but also students' learning process (Holstein et al. 2019b).

The glasses were tested using the ITS Lynnette, a linear equation tutor. Students were learning how to solve equations through Lynette and a teacher was monitoring them using Lumilo and providing assistance as needed. Lynette is a rule-based Cognitive Tutor, that has been shown to significantly improve students equation solving ability. It provides step by step guidance, gives hints when needed, and supplies feedback on the correctness. It also gives the students error-specific feedback messages while they solve the software's problems. Using Lynette the students face five levels of problems with increasing difficulty (Holstein et al. 2019b).

Lumilo glasses are transparent and allow the teacher to stay focused on the classroom but at the same time offer them support through analytics. There are three main types of displays: Student-level indicators, student-level „deep-dive“ screens and class-level summaries (Holstein et al. 2019b).

The student-level indicators and class-level summaries are visible by default, the student-level indicators can be seen above the student's heads, and the class-level summary is seen at a configurable location in the classroom. When glancing at a student indicator, a brief elaboration about the indicator symbol pops up. The indicator shows, for example, if a student is idle, tries to game the system(i.e. cheats and instead of learning just acts in a way to maximise his score) or has a high error rate after using a hint. By clicking on an indicator, the teacher can access the student's deep-dive screen, which includes a „Current Problem“ display, that shows a live feed of the students work on their current problem annotated with the number of hint requests and errors the student has made on that step. The deep dive also includes „Areas of Struggle“, the student's estimated probability of mastery, and concrete examples of error the student has made recently(Holstein et al. 2019b). At the class level overview the three skills that fewest of the students of the class mastered after attempting were shown. An example of the teacher displays can be seen in Figure 1. (Holstein et al. 2019b)

Using these analytics the teacher can keep an eye on the class, while directly intervening when a student needs help but is, for example, too shy to ask or the teacher knows when to help a student without even trying (Holstein et al. 2019b).

Figure 1: Holstein et al. (2019b) - Teacher’s point-of-view while using Lumilo. Top row: illustrative mock-ups; Bottom row: screenshots captured through Lumilo (taken after the end of a class session, to protect student privacy) [19]. Left: Teacher’s default view of the class through Lumilo. Right: Deep-dive screens that pop-up if a teacher ‘clicks’ on a student’s indicator .

Holstein tested the effectiveness of the tool in his paper on ”Co-designing areal-time classroom orchestration tool to support teacher”. Multiple classrooms were randomly assigned one of three conditions and analyzed. The conditions were „Glasses+Analytics“ where the teachers used the full Lumilo version including all the previously described features, „business-as-usual“ where they did not use the glasses, and only monitored the class the traditional way while they worked through the problems of the Lynette software and „Glasses“ where the teachers used a reduced version of Lumilo with solely the monitoring functionality. The last condition was added to make sure that the students’ mere awareness of the teacher’s monitoring had a significant effect on the student’s learning. In the experiment every class took a 20 minute computer-based pre-test, then worked with Lynette for 60 minutes and finally took a 20 minute computer-based post-test. During the pre- and post-test the students had no assistance from the teacher (Holstein et al. 2019b).

As can be seen in Figure 2 in all three conditions the students’ ability to solve equations increased. For students in the Glasses+Analytics condition, the post-test by pre-test curve however was comparatively flat, with lower pre-test students learning considerably more than in the other two conditions. There was no significant interaction between noGlasses/ Glasses and student pretest. However, there were significant negative interactions between student pretest scores and noGlasses/ Glasses+Analytics and Glasses/Glasses+Analytics. This suggests that using real-time analytics may be an equalizing force in the classroom, allowing teachers to allocate more time to students who actually need the help. This can also be seen in the bottom diagram of Figure 2 (Holstein et al. 2019b).

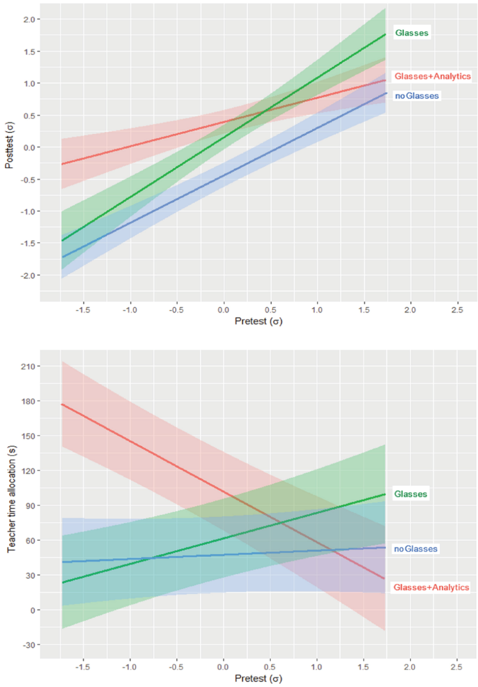


Figure 2: Holstein et al. (2019b) - Student post-test scores (top) and teacher attention allocation (bottom), plotted by student pretest scores, for each experimental condition. Shaded regions indicate standard error

This paints the picture, that learning analytics such as Lumilo are helpful for teachers and allow them to teach their students better, by granting them an insights to their learning behavior, allowing them to catch problems the students have, that might otherwise go unnoticed and intervene quickly. When the students' problems are caught early on this helps them to not fall behind. Lumilo also allows the teachers to effectively allocate their time to the students who especially need the support, and thus equalizes the performance in the classroom. Lumilo is thus an example of how Learning Analytics can and should be used to assist teachers, so they can teach their students the best they can.

4.2 Chatbots for language learning

Another tool that can be used as a teachers' assistant are chatbots. There are multiple ways a chatbot can be useful to a teacher. They can, for example, be used to answer students' simple problems. The chatbot functions here as an FAQ, that is available to the students 24/7, to explain organisational problems like simple parts of the syllabus or provide assignment's requirements like word count, deadline or other details, but also to answer simple questions about the topic, that is currently being studied. In this case the chatbot could also collect data on the questions the students ask, and problems they have, and could then give the teacher some further suggestions on learning material or problems that the students in general are struggling with. (Hamam 2021)

One of the main areas a chatbot can support a teacher is in language learning. In schools classes have often about thirty students, and with that many people it can become difficult to make sure, that every student gets to talk the same amount, and gets enough speaking, listening and comprehension practice. Chatbots can be helpful to teachers here, by providing a proper practicing environment, so that the teachers can focus more on teaching the language concepts like grammar. Students could, for example, engage with the chatbot in a foreign language in a natural manner to improve their reading and listening skills. This was the topic of a study by Lin & Mubarok (2021) on Learning Analytics for Investigating the Mind Map-Guided AI Chatbot Approach in an EFL Flipped Speaking Classroom from 2021. EFL in this case stands for English as a foreign Language. The flipped classroom is a pedagogical approach in which some activities (a task, homework or instructions) are swapped and the learning is taking place outside the classroom. The students learn new concepts at home and than consolidate their knowledge by doing training exercises in class. This is done to give the students a deeper learning experience than when the teacher guides them through the material(Lin & Mubarok 2021).

They used the Replika app, which is powered by artificial intelligence, to talk with humans via a chatbot. The students were using it to talk with the AI about a topic assigned by the instructor. The concept of the relationship between the robot and the student could be changed based on the need of the current conversation, such as friend, romantic partner or mentor, and the skills and traits of the chatbot could be changed as well. The chatbot also created tasks for the students, so that they would interact amongst them, such as sending videos, pictures or memes. The chatbot is shown in Figure 3 (Lin & Mubarok 2021).

The students would start by creating mindmaps using material provided by the teacher on the app, depending on their prior knowledge and level. Then they would practice with the chatbot based on the content of their mindmaps (Lin & Mubarok 2021). In the end students who interacted with the chatbot frequently, instead of doing low level activities such as doing a worksheet, spoke more fluently, used more accurate structures and developed the topics more coherently

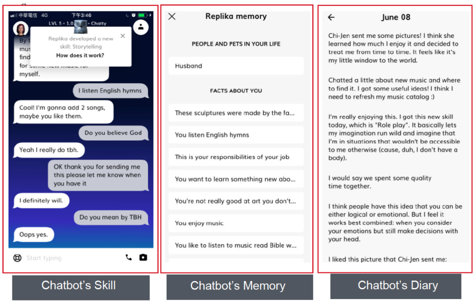


Figure 5. In-class and out-of-class learning activities using chatbots

Figure 3: Lin & Mubarok (2021) - Learning activities using the chatbot.

(Lin & Mubarok 2021). While an interaction with a teacher is still the most effective when learning a language, a chatbot could, for example, function as a home tutor, or help students learn the language who are to shy to talk to the teacher (Lin & Mubarok 2021).

In general chatbots can act as a cheap and easy tool for the students to interact with. They can provide the teacher with insights to the students' learning and provide a safe learning environment for students. They can function very well as an assistant to a regular human teacher, when used correctly.

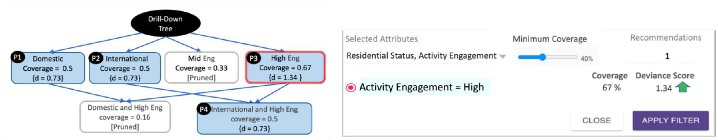
4.3 Learning analytics dashboard

Another tool aiding teachers both in and outside the classroom is the learning analytics dashboard. Defined by Schwendimann et al. (2016) as "a single display that aggregates different indicators about learner(s), learning process(es) and/or learning context(s) into one or multiple visualizations", learning analytics dashboards are a tool used to visualize the wealth of information gathered on learning behavior. Their application is diverse ranging from supporting traditional face-to-face lectures or group-work to aiding in online and blended learning approaches. Generally the dashboards provide feedback and prognoses enabling both teachers and students to make informed decisions on the learning progress and further learning process (Klerkx et al. 2017).

In their review Schwendimann et al. 2016 found that dashboards are currently mainly used in formal learning in higher education if the educational level is at all specified. This is a rather narrow field so the application of learning analytic dashboards should be explored in a wider context. Another issue is that visualization closely resembles other areas of dashboard applications like bar plots or similar diagrams. Taken together with the growing complexity of the data the learning analytics dashboards are supposed to represent, this highlights the lack of learning specific visualizations.

Still, learning analytics dashboards can greatly enrich the teacher's knowledge about the progress and learning behavior of their students and can aid them in anticipating a student's success or need for aid. As such it is important to gather large quantities of data on the student's learning behavior. This is usually done by tracking the use of online modules, learning artifacts, questionnaires, assessments and more. Dashboards implemented for classroom settings may use a multimodal approach tracking movement, gaze gestures, facial expression and interactions besides the traditional assessment factors as well. To enable this, learning analytics dashboards usually appear in computer-supported learning.

The dashboards usually rely heavily on predictive learning analytics. This is a problem since the data the dashboards are supposed to represent is growing increasingly complex. Khosravi et al. (2021) introduce an approach called automated insightful drill-down (AID) to switch from a predictive learning approach to a recommendation one. Drill-down operations progressively apply filters to the input data enabling a more granular analysis. AID first drills down in the data gathered on the student's study behavior to create a tree, which leads to sub-populations within the students, based on predefined filter categories, like specific behavior or social data. This tree gets pruned by erasing all nodes with a sub-population smaller than a given threshold. In a second step the distance to the average population is calculated. In this way a teacher can easily find outlier populations, which may need more attention than the average student population (see Figure 4a). To simplify the search for specific sub-populations, the learning analytics dashboard includes a filter function and a manually entered coverage (see Figure 4b).



(a) Drill-down tree for performance-based recommendations. (b) An example drill-down recommendation interface.

Figure 5. The drill-down tree and the associated drill-down recommendation for the performance-based illustrative example.

Figure 4: Khosravi et al. (2021: p. 9) - Drill-down recommendation and interface visualization.

Besides identifying populations and students which require aid, it si especially important for teachers to get feedback from the students. This can be difficult during e-learning, since the teacher lacks most of the usual visual and verbal cues from the students in an online setting. These cues usually allow them to react to their class and tailor lessons to the current needs of the students. Franzoni et al. (2020) proposed a tool embedded in a Learning Management System (LMS) that visualizes a student's engagement in color and dimensions of learning objects and their individual avatars. They implemented this visualization as a plug-in for the activity logs of one course on the learning platform Moodle and found their visualization techniques worked better for the teachers in tracking overall engagement than the standard Moodle visualization.

In general it can be said that learning analytics dashboards are an easy to use tool that aids teachers and students alike in optimizing their performance based on immediate feedback. Furthermore such dashboards can be used to generate prognoses based on pre-existing aggregated data that has been gathered in comparable situations for the current application. But while learning analytics dashboards seem to be a worthwhile endeavor to visualize highly complex data, it is still work in progress and needs to be adapted to the users' requirements to enable efficiency and impact. Moreover, it progresses to be a more and more complex task to represent the growing multidimensional data we feed into our learning analytics, without either overcrowding the dashboard or loosing important information. Finally, it is important to integrate the tracked learning progress and predicted success in the context of academic achievements. While one can track usage behavior and make predictions based on them, regular tests and exams are still needed to make sure the student actually learned the content.

5 Conclusion

Taking all these Application, as well as many others we have not touched upon into consideration, it is clear to see, how widespread and fast growing this field is. However, even if there is a lot of work done to create smart tools to assist teachers in and out of the classroom, most tools have not progressed far enough to be generally usable. All the tools discussed in this paper, as different as they are, are all in the beginning stages of development. Nevertheless, the speed with which these first examples were developed, does give hope for what tools could emerge and mature in the next couple of years, that can help teachers improve on their teaching. Keeping this in mind, it definitely shows that this field of AI support for teachers is definitely worth pursuing further.

References

Franzoni, Valentina, Alfredo Milani, Paolo Mengoni & Fabrizio Piccinato. 2020. Artificial intelligence visual metaphors in E-Learning interfaces for learning analytics. *Applied Sciences* 10(20). 7195. DOI: 10.3390/app10207195.

Hamam, Doaa. 2021. The New Teacher Assistant: A Review of Chatbots' Use in Higher Education. *Communications in Computer and Information Science* 1421. 59–63. DOI: 10.1007/978-3-030-78645-8_8.

Holstein, Kenneth, Bruce M McLaren & Vincent Aleven. 2019a. Co-designing a real-time classroom orchestration tool to support teacher - AI complementarity. *Journal of Learning Analytics* 6(2). 27–52. DOI: 10.18608/jla.2019.62.3.

Holstein, Kenneth, Bruce M McLaren & Vincent Aleven. 2019b. Designing for complementarity: Teacher and student needs for orchestration support in ai-enhanced classrooms. 157–171. DOI: 10.1007/978-3-030-23204-7_14.

Hsieh, YiZeng, ShihSyun Lin, YuCin Luo, YuLin Jeng, ShihWei Tan, Chao-Rong Chen & PeiYing Chiang. 2020. ARCS-Assisted Teaching Robots Based on Anticipatory Computing and Emotional Big Data for Improving Sustainable Learning Efficiency and Motivation. *Sustainability* 12(14). 5605. DOI: 10.3390/su12145605.

Khosravi, Hassan, Shiva Shabaninejad, Aneesha Bakharia, Shazia Sadiq, Marta Indulska & Dragan Gašević. 2021. Intelligent Learning Analytics Dashboards: automated Drill-Down Recommendations to Support Teacher Data Exploration. *Journal of Learning Analytics* 8(3). 133–154. DOI: 10.18608/jla.2021.7279.

Klerkx, Joris, Katrien Verbert & Erik Duval. 2017. *Chapter 12: Learning analytics dashboards.* Society for Learning Analytics Research. DOI: 10.18608/hla17.012.

Lin, Chi-Jen & Husni Mubarok. 2021. Learning Analytics for Investigating the Mind Map-Guided AI Chatbot Approach in an EFL Flipped Speaking Classroom. *Educational Technology and Society* 24(4). 16–35.

Majumdar, Rwitajit, Brendan Flanagan & Hiroaki Ogata. 2021. eBook Technology Facilitating University Education During COVID-19: Japanese Experience. *Canadian Journal of Learning and Technology* 47(4). DOI: 10.21432/cjlt28038.

Michos, Konstantinos, Charles Lang, Davinia Hernández-Leo & Detra Price-Dennis. 2020. Involving teachers in learning analytics design: Lessons learned from two case studies. 94–99. DOI: 10.1145/3375462.3375507.

Rincón-Flores, Elvira G, Eunice López-Camacho, Juanjo Mena & Omar Olmos López. 2020. Predicting academic performance with Artificial Intelligence (AI), a new tool for teachers and students. 2020. 1049–1054. DOI: 10.1109/EDUCON45650.2020.9125141.

Schwendimann, Beat A, Maria Jesus Rodriguez-Triana, Andrii Vozniuk, Luis P Prieto, Mina Shirvani Boroujeni, Adrian Holzer, Denis Gillet & Pierre Dillenbourg. 2016. Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies* 10(1). 30–41. DOI: 10.1109/TLT.2016.2599522.

Chapter 6

AI-based adaptive gamification in math learning scenarios

Pia Tamina Ondreka, Carlos Parra Fernandez, Frederik Wollatz & Qirui Zhu

In math learning scenarios, learners are often facing motivational and emotional issues due to the perceived high level of difficulty. However, adaptive learning strategies that use Artificial Intelligence (AI) or gamification try to counteract these problems. In this literature review we reviewed 25 publications to determine the current progress on gamification, AI and math learning. Each of the three topics themselves are well researched, but cross-collaborative work remains sparse. While gamification is suitable for addressing motivational issues, i.e., by using game mechanics like challenges, goals, story-telling or rewards, AI is a common choice for generating personalized content due to its ability to adapt and thus is applicable in terms of gamified approaches. In addition to gamification, Intelligent Tutoring Systems (ITS) are also widely used in math learning, but the combination of both is rare. However, combining these two approaches could benefit from both advantages of improving motivation while offering individualized education to enhance the learning experience.

Keywords: Artificial Intelligence, Machine Learning, Gamification, Education, E-Learning, Math Learning

1 Introduction

In recent years, the field of Artificial Intelligence (AI) in Education gave rise to new technical approaches in the field of traditional education. Simultaneously, the amount of publications concerning the usage of game-like elements in an educational setting increased, establishing the field of gamification.

Gamification is defined as the usage of game mechanics in a non-game context or experiencing a game experience in serious contexts (Krath et al. 2021), i.e., traditional educational contexts. In contrast, the notion of game-based learning describes the process of learning with serious games, developed for the purpose of education (Krath et al. 2021) with a more short-termed focus (Duggal et al. 2021). As teachers and educators are facing more motivational issues among their students (Tanjung & Sitompul 2020), the implementation of gamification can facilitate the creation of immersive learning environments for self-regulation and development of intrinsic motivation and control (Bertram 2020). Among the most used common game mechanics are reward systems, badges, leader-boards and virtual currencies (Koravuna & Surepally 2020). This paper will mainly focus on digital gamification and game-based learning due to the nature given by the AI field. Despite the aforementioned differentiation of game-based learning and gamification in the literature, we will refer to the gamification term as a broader term covering both of these notions. More gamification techniques are discussed in the sections 3.1, 3.2 and 3.3.

Artificial Intelligence approaches, including machine learning, have already established themselves in the educational field as techniques for analyzing student performance and detecting disengagement (Duggal et al. 2021). Syzdykbayeva et al. (2021) describe AI usage in higher education under the current pandemic circumstances, highlighting the need for personalized and adaptive learning content. As education in general is moving towards Massive Open Online Courses (Duggal et al. 2021), providing individualized learning support becomes less feasible with our current technology, thus making AI a promising solution for providing personalized content. We will continue to discuss some AI and machine learning techniques in the context of gamification in section 3.2.

Mathematics is still one of the most important subjects especially in technology-heavy fields. One major problem of learning mathematics is that many students experience difficulties with the subject, leading to demotivation and dropout (Rahim et al. 2019). Hussein et al. (2021) report that math learning often causes frustration, fatigue, pressure and anxiety among the students which results in high failure rates. The student´s emotions are especially important in math learning scenarios, which are influenced by the controllability of the learning situation and their expectations, attributes and competency beliefs (Bertram 2020). In terms of gamified approaches to math learning Bouzid et al. (2021) found that the traditional paper-sessions are more deteriorating than game-based sessions in terms of math anxiety (MA) which is a big factor in math learning for students. In addition Bouzid et al. (2021) also found that the attitudes towards in-class mathematics activities positively shifted after game-sessions, indicating that gamification can help easing MA and improve classroom experience. More math learning scenarios and the usage of gamification in these scenarios are presented in section 3.1 while section 3.3 deals with the combination of gamification, math learning and AI.

In order to give an overview of the current theoretical and practical concepts of machine learning and AI techniques in combination with gamification elements in the educational setting of math learning scenarios, we conducted a literature review with the following main research questions.

- How is gamification realized in mathematics learning?
- How is AI used in combination with gamification?
- How can the concepts of AI, gamification and math learning be combined interdisciplinary to facilitate cross-domain collaboration?

2 Methodology

This literature review was conducted by screening papers from Elsevier´s Scopus and Google Scholar databases. According to the three main topics of gamification, math learning and artificial intelligence several search strings were used. All the used keywords can be found in table 1. The search strings consist of the combination of search terms from two of the main topics or from all three topics, e.g., for the topic gamification in AI a corresponding search string was ”gamification AND machine learning” or ”gamification AND ai OR machine learning”.

Table 1: Search strings

Topic	Search terms
Gamification	”gamification” OR ”game” OR ”adaptive AND gamification”
Artificial Intelligence	”ai” OR ”artificial intelligence” OR ”machine learning”
Math learning	”math learning” OR ”math” OR ”math AND education”

For the Scopus database, the search was limited to the title, abstract and keywords. For all the entries found in the first pages of both databases, the abstracts were evaluated according to the existence of the search terms and the applicability of the paper. For instance, papers were excluded which contained the keyword ”artificial intelligence” and ”gamification”, but dealt with the teaching of ”artificial intelligence” as a subject, due to the lack of actual AI techniques being used. Additionally all corresponding journals were checked in terms of peer-review or to-be-peer-reviewed (e.g., arxiv.org) status, as well as their status of being a potential predatory journal. Following this procedure the literature acquisition yielded a total of 25 references, constituting this literature review.

3 Results

While AI is a widespread topic in gamification and education, we could only find one study (Faghihi et al. 2017) that deals with AI-based gamification in a math learning context. However, even Faghihi et al. (2017) hardly make any connection between AI in gamification and gamification in math learning. Therefore, we want to focus on the individual subject areas first and then make a suggestion for a possible combination.

3.1 Gamification and math learning

Digital games are generally fun and encourage active engagement. Combined with the learning process, this can increase intrinsic motivation, which is one of the key aspects of gamification of learning (Yong et al. 2021). But this also means that gamification has to address the basic needs of a learner so that motivation is maintained (Krath et al. 2021). Autonomy, competence, relatedness and self-efficacy can be influenced in a positive manner by different gamification methods (Krath et al. 2021). In addition, gamification also has positive effects on learning behavior (e.g., engagement), cognitive skills (e.g., problem solving) and knowledge acquisition (Hussein et al. 2021, Krath et al. 2021).

These positive effects result from the large number of different game mechanics that can specifically target the individual needs of a learner. Some common game mechanics are challenges, competition, feedback, rewards, story-telling and goals (Duggal et al. 2021, Faghihi et al. 2014, 2017, Krath et al. 2021, Yong et al. 2021). In math learning games, game mechanics such as story-telling, collecting coins and guidance are already used (Rahim et al. 2019).

Challenges can be used to address a learner’s need of competence and self-efficacy (Yong et al. 2021). However, they have to be constructed in such a way that they are feasible in principle, albeit with great effort, because easily achievable challenges can lead to boredom, while unachievable challenges can trigger stress and anxiety (Yong et al. 2021). With an increasing level of difficulty, challenges can be adapted to the skills and knowledge of the learner, which also results in better self-efficacy (Krath et al. 2021). Combined with direct feedback and the possibility of success or failure, the learner should be encouraged to invest more effort in the learning process without becoming frustrated (Yong et al. 2021). This usually leads to fun with the challenge itself, i.e., process-focused learning and not just with the goal achievement (Yong et al. 2021).

Immediate feedback also contributes to a better feeling of competence (Krath et al. 2021). Especially in math learning, a mistake in long calculation chains is often only noticed at the end, leading to an inefficient and time-consuming process of troubleshooting and revision (Yong et al. 2021). To prevent this, the calculation can be divided into smaller steps with instant feedback, for example using levels, point systems or progress bars (Krath et al. 2021). Long-term feedback in the form of positive reinforcement through badges, achievements, gifts or rewards also promote competence and self-efficacy and contribute to greater engagement and satisfaction (Krath et al. 2021). On the other hand, these mostly improve the extrinsic motivation (i.e., performing a behavior for an expected reward) rather than intrinsic motivation (i.e., performing a behavior for its own sake) and thus, according to Zainuddin et al. (2020), should not be overused.

In order to achieve learning success, clear and relevant goals should be set that correspond to the needs of the learner according to specificity and difficulty so that the learner remains motivated (Krath et al. 2021). Math problems and, accordingly, the formulation of goals, should be comparable to problems in the real world, so that the experience gained can contribute to the learning process (Yong et al. 2021). Learners should also be able to set own goals in order to support self-regulated learning and thus autonomy. This can be implemented by leader-boards, for example (Krath et al. 2021). Goal systems can be supported by challenges, level systems, quests or stories (Krath et al. 2021). Story-telling can help guide the learner on the way to the learning goal (Krath et al. 2021) and gives meaning to the learning process by making the learner part of the story (Yong et al. 2021). Within the story, the learner can encounter realistic situations in the virtual environment, which supports learning from experience and also makes learning more interesting and fun (Yong et al. 2021). Nevertheless, the possibility to find individual ways to achieve a goal should remain open, for example with non-linear game-play in order to strengthen the learner's autonomy (Krath et al. 2021, Faghihi et al. 2017). Customization in general (Krath et al. 2021), but also opportunities for exploration, testing different strategies, self-expression and discovery help satisfying the need for autonomy as well (Yong et al. 2021).

In order to meet the need for relatedness, elements can be used that promote collaboration or competition (Krath et al. 2021, Yong et al. 2021). Above all, competition plays a major role in motivation, engagement and learning performance (Zainuddin et al. 2020), especially for learners with a poor socio-economic background (Hussein et al. 2021). Leader-boards work best here (Zainuddin et al. 2020), but duels or contests are also possible implementations (Krath et al. 2021). On the other hand, there should also be opportunities for learners to support each other and work together in groups. Means of communication and team challenges can be provided for this (Krath et al. 2021). Games are able to create new kinds of relationships between people like friendship or animosity simply by telling a story in a role-play scenario (Yong et al. 2021).

In addition to individual game mechanics, adaptive learning environments and Intelligent Tutoring Systems (ITS) are particularly suitable for gamified learning processes (Faghihi et al. 2014, 2017). ITS aim to motivate learners by providing information when necessary. However, they lack fun and enjoyment, which is why the intrinsic motivation to use such a system may be lacking as well (Faghihi et al. 2014). Adaptive learning systems, on the other hand, adapt to the cognitive, motivational and emotional states of the learners and are therefore better suited to meeting their needs (Bertram 2020). Combining ITS and adaptive gamification can involve, for example, (1) feedback and competition by seeing the individual and others' levels, (2) receiving hints when they are needed, (3) rewards for better motivation, (4) the possibility of level-up when reaching a goal (Faghihi et al. 2014) or (5) adjustments to design, instructions or game content (Bertram 2020).

When it comes to math, there are already many games, most of which deal with the subjects of operations, algebra, geometry, measurements and data analysis and probabilities (Bertram 2020). The majority of these games are educational games and therefore designed for this specific subject. However, games not specifically developed for education can also be used to gain an intuition on certain topics (Yong et al. 2021). In general, games for learning linear algebra are highly accepted by learners (Rahim et al. 2019). Rahim et al. (2019) were able to collect some essential game mechanics for this acceptance, including reward collection, feedback, story-telling, enhancement of the knowledge-base, time pressure, the need to work harder and unexpected surprises and losses. They also recommend a reward system with points, different levels of difficulty and badges and achievements, which are presented on a leader-board, as well as goals and challenges, including the opportunity for cooperation (Rahim et al. 2019). Faghihi et al. (2017) have developed an algebra game that combines ITS and adaptive gamification and uses some of these game mechanics. It provides a story, different levels of difficulty, information about how the game works and how the underlying math problem can be solved, scores, hints and adapts the design and speed according to the learners' needs. Using ITS and gamification improves learning efficiency and learners are shown to prefer the game elements to pure ITS because they are more fun (Faghihi et al. 2014). In addition to classic games, interactive math e-books can also be used in a gamified manner (Zhao et al. 2021). Overall, it has been shown that playing math games reduces MA (Faghihi et al. 2017), contributes to more relaxed, motivated and comfortable learning and significantly improves the learning process (Barbieri et al. 2021).

3.2 Gamification and AI

Since the use of Artificial Intelligence in education is slowly rising, the use of Gamification to support these tools is increasing as well. In the last few years, AI has caught attention due to its ability to provide autonomous and adaptive functions usable for offering personalized education (Bezzina et al. 2021). According to Khakpour & Colomo-Palacios (2021), learning, personalization and behavioral change are among the top five fields that have an interest in pushing these technologies. Even then, Khakpour & Colomo-Palacios 2021 point out that gamified tasks often do not remain of long-term interest to the user.

The typical application is the embedding of these gamifications into an already existing learning platform (Khakpour & Colomo-Palacios 2021). To get the most out of gamification, AI can be used to choose personalized gamification elements for the learner. These systems capture the interaction between system and user in order to identify and adapt to skills, preferences and affective states (Monterrat et al. 2014, Bertram 2020). Three different data sources can be used: user data (including age and personality traits), usage data (e.g., login times) and environment data (e.g., where the game is played) (Monterrat et al. 2014). This data can be collected before, but also during the use of gamification (Bertram 2020).

Monterrat et al. (2014) suggested separating player model from learner model. The player model should focus on adaptation while the learner model takes care of educational needs. In addition, Knutas et al. (2017) described a proof-of-concept algorithm that poses challenges for different user types. These user types are based on the Gamification User Types Hexad Scale (Tondello et al. 2016), which categorizes users into five groups: Philanthropist, Socialiser, Free Spirit, Achiever and Player. Based on the player type and the appropriate tasks, the system provides the user with quests that match the player's interest, thus increasing the long-term interest in learning. A similar grouping of players was also performed by Daghestani et al. (2020), separating the user into Conquerors, Socialiser, Achiever, Mastermind and Seeker. Based on these profiles, they applied gamification techniques according to Ferro et al. (2013). Results showed significantly higher engagement and knowledge test scores for the adaptive gamification group compared to a normal gamification group and a non-gamification group (Daghestani et al. 2020).

When using a learning game for the first time, however, there is the problem that the exact learner model is not yet known. Accordingly, a good learning experience cannot be guaranteed from the system side, especially with ITS, which is only reinforced on the user side by the unfamiliarity of the new learning environment (Pian et al. 2020). This so-called cold-start problem can be tackled by the Mechanics-Dynamics-Aesthetics (MDA) game design framework, which introduces the system to the user while collecting data at the same time (Pian et al. 2020). The MDA framework includes six game mechanics, separated into narrative elements (i.e., background story, system avatar and role playing) and task-oriented elements (i.e., missions, rewards and feedback) for data collection. This makes an acquisition of 92,5% of the necessary user data possible and results in a higher learner engagement and usability (Pian et al. 2020).

Bertram (2020) proposed to use Optimal Experimental Design (OED), developed for optimizing experimental designs according to various psychological models. In OED, the interactions between system and human are small experiments, the results of which are used to refine the user profile. It provides the user with situations that are most promising for creating a learner model based on performance, motivation, engagement and emotional state (Bertram 2020).

Finally, Hassan et al. (2021) suggest basing the adaptation on learning styles, following Felder & Silverman (1988). In order to identify the learning styles, the interaction times with the various modules, e.g., the gamification, collaboration and content modules of the system are logged. This separates the learners along the four dimensions <Active, Reflective>, <Sensing, Intuitive>, <Sequential, Global> and <Visual, Verbal> (Felder & Silverman 1988). Testing one group with adaptive gamification and one with fixed gamification showed a significant increase in the course completion rate for the group with adaptive gamification. In addition, motivation and the amount of interactions with the system increased significantly (Hassan et al. 2021).

However, even if several AI assisted systems have been developed in recent years, there is still a lack of scientific evidence regarding the impact of these systems on the overall learning quality (Bezzina et al. 2021).

3.3 Combining AI, gamification and math learning

While gamification in math learning has shown to have many advantages, there are still some latent problems that need to be addressed in order for the effectiveness of this type of learning to be improved. More specifically, given a particular class the skill level of the individual students can be very widespread. Due to this it is not possible to provide individual support for every single learner. To solve this problem, ITS can be combined with gaming technologies to individualize the learning process, but for that the system needs to be adaptive (Faghihi et al. 2014). While the ITS is responsible for adapting the math learning context, the adaptive gamification system takes care of the gaming context. Both are controlled by the responsible AI. This ensures that the difficulty of the current and subsequent exercises is appropriate for each learner, but also that the respective rewards, feedback, challenges and points are attributed (Faghihi et al. 2017). The AI's decisions and contributions would replace or simulate those that teachers make in such scenarios. Particularly, the teachers' contributions in the study of Barbieri et al. (2021) have shown to be decisive in changing game mechanics to appropriately adapt them to the students' training needs in math learning.

Studies like the one performed by Zhao et al. (2021) show that providing students with a gamified math learning platform where they can share their progress and communicate with peers, allowed them to better support each other and increase their engagement within the environment. This could be adapted using machine learning to allow the ITS to better pair classmates on the networking platforms, included as game mechanic, so that these are able to help each other out more optimally and dynamically (Krath et al. 2021).

Enhancing of gamified learning engagement together with AI is a technique present in online games and training simulators for AI and even language learning applications (Syzdykbayeva et al. 2021), but has yet to be implemented in widespread math learning applications. However, if carried out effectively it could potentially deal with the present MA, allowing the knowledge and pressure to much more smoothly ease off and blend into the learning environments of affected learners. This adaptability of gamification needs to be acquired, so the correct balance in the mechanics is implemented. If the feedback provided through gamification fails to attain proper equilibrium this will work as a double-edged sword, which can influence the learning experience and effectiveness negatively (Bouzid et al. 2021) e.g., learners completing tasks just for the sake of earning an achievement and not actually for the intrinsic desire of learning further. This makes the gamified learning into just a game, where the player is not actually learning, but rather finishing off tasks without much thought.

4 Conclusion

Gamified learning and serious games seem to provide many advantages to the learners. For some learners suffering from MA, gamified learning provides a more approachable environment in which learners can tackle the subject from a more comfortable stance in comparison to the usual paper-sessions and after game-sessions activities related to mathematics are perceived more positively by the learners (Bouzid et al. 2021). Gamification could also manage to counter the struggle of some learners' lack of motivation, caused by the lack of ability to keep alert during class or maybe even due to psychological disorders (Duggal et al. 2021) (see chapter IV for more examples), especially those attached to math learning itself.

Gamification should aim to provide a proper overlap between technical knowledge, content knowledge and pedagogical knowledge (Barbieri et al. 2021). For this, the limitations of digital learning need to be taken into account, since some intrinsic features of the embodied learning either cannot be replaced by the digital learning or are too costly in order to be effective in this context, particularly the discrimination against learners with no easy access to the proper technology, a lack of familiarity with the environment and even the risks tied to the data-security complications that could arise (Bertram 2020). Once these obstacles have been overcome, actual analysis of individual learners' performance and their engagement can be better measured, monitored and adjusted using the established technique of machine learning (Duggal et al. 2021). This is key so as to regulate the amount of feedback provided by the gamification methods implemented. If these were not to be balanced properly, it could damage the learners' intrinsic motivation (Zainuddin et al. 2020). With proper adaptability of the ITS implemented even learners that were already performing relatively well at math might be able to actually refine their skills through this adaptations, since they did not show any particular im-provements with non-adaptive learning games (Barbieri et al. 2021).

Overall, the fields of AI, gamification and math learning research have each advanced significantly over the recent years, but there is still a gap to be filled with further research, covering the combination of those topics to provide an overview on the practical effectiveness.

References

Barbieri, Giuseppina Gerarda, Rosa Barbieri & Roberto Capone. 2021. Serious games in high school mathematics lessons: an embedded case study in europe. *EURASIA Journal of Mathematics, Science and Technology Education* 17(5). em1963. DOI: 10.29333/ejmste/10857.

Bertram, Lara. 2020. Digital learning games for mathematics and computer science education: the need for preregistered RCTs, standardized methodology, and advanced technology. *Frontiers in Psychology* 11. 2127. DOI: 10.3389/fpsyg.2020.02127.

Bezzina, Stephen, Alexander Pfeiffer & Alexiei Dingli. 2021. AI-enabled gamification for learning and assessment. In *8th International Conference on Educational Technologies 2021, ICEduTech 2021 and 17th International Conference on Mobile Learning 2021, ML 2021*, 189–193. Virtual event: IADIS Press.

Bouzid, Tariq, Fatiha Kaddari, Hassane Darhmaoui & El Ghazi Bouzid. 2021. Enhancing math-class experience throughout digital game-based learning, the case of moroccan elementary public schools. *International Journal of Modern Education and Computer Science* 13(5). 1–13. DOI: 10.5815/ijmecs.2021.05.01.

Daghestani, Lamya F., Lamiaa F. Ibrahim, Reem S. Al-Towirgi & Hesham A. Salman. 2020. Adapting gamified learning systems using educational data mining techniques. *Computer Applications in Engineering Education* 28(3). 568–589. DOI: 10.1002/cae.22227.

Duggal, Kavisha, Lovi Raj Gupta & Parminder Singh. 2021. Gamification and machine learning inspired approach for classroom engagement and learning. *Mathematical Problems in Engineering* 2021(9922775). 1–18. DOI: 10.1155/2021/9922775.

Faghihi, Usef, Donald Aguilar, David Chatman, Nicholas Gautier, Jeffrey Gholson, Justin Gholson, Melvin Lipka, Robert Dill, Philippe Fournier-Viger & Sioui Maldonado-Bouchard. 2017. How to apply gamification techniques to design a gaming environment for algebra concepts. In *E-Learning, E-Education, and Online Training*, vol. 180 (Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering), 62–70. Dublin, Ireland: Springer, Cham. DOI: 10.1007/978-3-319-49625-2\{_}8.

Faghihi, Usef, Albert Brautigam, Kris Jorgenson, David Martin, Angela Brown, Elizabeth Measures & Sioui Maldonado-Bouchard. 2014. How gamification applies for educational purpose specially with college algebra. *Procedia Computer Science* 41. 182–187. DOI: 10.1016/j.procs.2014.11.102.

Felder, Richard M. & Linda K. Silverman. 1988. Learning and teaching styles in engineering education. *Engineering Education* 78(7). 674–681.

Ferro, Lauren S., Steffen P. Walz & Stefan Greuter. 2013. Towards personalised, gamified systems: an investigation into game design, personality and player typologies. In *Proceedings of The 9th Australasian Conference on Interactive En-*

tertainment: Matters of Life and Death (IE '13), 1–6. Melbourne, Australia: Association for Computing Machinery. DOI: 10.1145/2513002.2513024.

Hassan, Muhammad Awais, Ume Habiba, Fiaz Majeed & Muhammad Shoaib. 2021. Adaptive gamification in e-learning based on students' learning styles. *Interactive Learning Environments* 29(4). 545–565. DOI: 10.1080/10494820.2019.1588745.

Hussein, Mahmood H., Siew H. Ow, Monther M. Elaish & Erik O. Jensen. 2021. Digital game-based learning in K-12 mathematics education: a systematic literature review. *Education and Information Technologies*. DOI: 10.1007/s10639-021-10721-x.

Khakpour, Alireza & Ricardo Colomo-Palacios. 2021. Convergence of gamification and machine learning: a systematic literature review. *Technology, Knowledge and Learning* 26(3). 597–636. DOI: 10.1007/s10758-020-09456-4.

Knutas, Antti, Rob van Roy, Timo Hynninen, Marco Granato, Jussi Kasurinen & Jouni Ikonen. 2017. Profile-based algorithm for personalized gamification in computer-supported collaborative learning environments. In *Proceedings of the 1st Workshop on Games-Human Interaction*, vol. 1956 (GHITALY 2017), 1–6. Cagliari, Italy: CEUR Workshop Proceedings. DOI: 10.5281/zenodo.827225.

Koravuna, Shamini & Uday Kumar Surepally. 2020. Educational gamification and Artificial Intelligence for promoting digital literacy. In *Proceedings of the 2nd International Conference on Intelligent and Innovative Computing Applications* (ICONIC '20), 1–6. Plaine Magnien Mauritius: Association for Computing Machinery. DOI: 10.1145/3415088.3415107.

Krath, Jeanine, Linda Schürmann & Harald F. O. von Korflesch. 2021. Revealing the theoretical basis of gamification: a systematic review and analysis of theory in research on gamification, serious games and game-based learning. *Computers in Human Behavior* 125. 106963. DOI: 10.1016/j.chb.2021.106963.

Monterrat, Baptiste, Elise Lavoué & Sébastien George. 2014. Motivation for learning: adaptive gamification for web-based learning environments. In *Proceedings of the 6th International Conference on Computer Supported Education* (CSEDU-2014), 117–125. Barcelona, Spain: SCITEPRESS - Science & Technology Publications, Lda. Av. D. DOI: 10.5220/0004848101170125.

Pian, Yang, Yu Lu, Yuqi Huang & Ig Ibert Bittencourt. 2020. A gamified solution to the cold-start problem of intelligent tutoring system. In *Artificial Intelligence in Education: 21st International Conference, AIED 2020*, vol. 2 (AIED 2020), 376–381. Ifrane, Morocco: Springer-Verlag, Berlin, Heidelberg. DOI: 10.1007/978-3-030-52240-7\{_}68.

Rahim, Rini H. A., Aslina Baharum & Hanafi Hijazi. 2019. Evaluation on effectiveness of learning linear algebra using gamification. *Indonesian Journal of Electrical Engineering and Computer Science* 17(2). 997–1004. DOI: 10.11591/ijeecs.v17.i2.pp997-1004.

Syzdykbayeva, Aigul, Aigerim Baikulova & Rysty Kerimbayeva. 2021. Introduction of Artificial Intelligence as the basis of modern online education on the example of higher education. In *2021 IEEE International Conference on Smart Information Systems and Technologies* (SIST), 1–8. Nur-Sultan, Kazakhstan: IEEE. DOI: 10.1109/SIST50301.2021.9465974.

Tanjung, Mahardika Abdi Prawira & Opim Salim Sitompul. 2020. "Kian Santang" game as historical educational media using digital storytelling concept. *Education and Information Technologies* 25(2). 5379–5388. DOI: 10.1007/s10639-020-10209-0.

Tondello, Gustavo F., Rina R. Wehbe, Lisa Diamond, Marc Busch, Andrzej Marczewski & Lennart E. Nacke. 2016. The gamification user types hexad scale. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play* (CHI PLAY '16), 229–243. Austin, Texas, USA: Association for Computing Machinery. DOI: 10.1145/2967934.2968082.

Yong, Su-Ting, Natanael Karjanto, Peter Gates, Tak-Yee A. Chan & Than-Myint Khin. 2021. Let us rethink how to teach mathematics using gaming principles. *International Journal of Mathematical Education in Science and Technology* 52(8). 1175–1194. DOI: 10.1080/0020739X.2020.1744754.

Zainuddin, Zamzami, Samuel Kai Wah Chu, Muhammad Shujahat & Corinne Jacqueline Perera. 2020. The impact of gamification on learning and instruction: a systematic review of empirical evidence. *Educational Research Review* 30. 100326. DOI: 10.1016/j.edurev.2020.100326.

Zhao, Jiahua, Gwo-Jen Hwang, Shao-Chen Chang, Qi-fan Yang & Artorn Nokkaew. 2021. Effects of gamified interactive e-books on students' flipped learning performance, motivation, and meta-cognition tendency in a mathematics course. *Educational Technology Research and Development* 69(3). 3255–3280. DOI: 10.1007/s11423-021-10053-0.

Part II

Methods and Technologies

Chapter 7

Comparing different spaced repetition algorithms for enhancing human learning

Marko Duda, Niklas Dettmer

Spaced repetition is a learning method for memorizing items and is well known since the 70s. With the raise of modern AI and learning platforms like Duolingo, a lot of new algorithms were introduced for optimizing spaced repetition. This chapter provides an overview of different classes of spaced repetition algorithms by using a selection of the most significant papers in the field and shows how they perform in terms of recall and activity rate. Considering a productive setting, different conceptual strengths and weaknesses of these algorithms are presented. We show that both, a rule-based as well as a ruleless algorithm are interesting for optimizing spaced repetition.

Keywords: AI, Deep Reinforcement Learning, Spaced Repetition, Duolingo, Memorization, Optimization, Vocabulary

1 Introduction

Spaced repetition is a technique for effective scheduling of items that need to be memorized by a learner. An item can be anything that can be memorized like a vocabulary or flash cards. Spaced repetition makes use of two psychological effects: first, the spacing effect says that short study practices spread over time are more effective than cramming in a short time period (Ebbinghaus 1885) when it comes to memorization. Second, the lag effect (Melton 1970) states that people learn better if the spacing between memorizing an item gradually increases. Both effects are well documented when it comes to second language acquisition (Settles & Meeder 2016). The most basic algorithms for spaced repetition are quite simple, as they do not require AI or even a computer and can be applied by the learner directly. Take the Leitner system for example: it uses different boxes (see Figure 1) which correspond to different practice intervals (1-day, 2-day, 3-day and so on). All items start on the 1-day interval box. When a vocabulary is learned, it moves one box further to the right and when it is not recalled correctly by the learner, it moves a box back to the left.

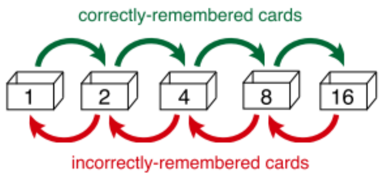


Figure 1: The Leitner system for spaced repetition

However, the Leitner system is very static. When a newly introduced item is already known by the student before, it will be queried multiple times nonetheless. Modern AI enables spaced repetition scheduling that is not as static as the Leitner system. Furthermore, modern second language acquisition platforms like Duolingo (they call themselves the most popular platform for that purpose[1]) gather massive amounts of user data which can even be used to make individual learning profiles of each student. When applied in these AI systems, spaced repetition aims for two goals:

- Memorizing an item as efficient as possible: items should not be queried more often than necessary for retaining in long-term memory.
- Making the learning experience pleasant: students should be encouraged to use an App like Duolingo more often.

Both of these goals can be evaluated quantitatively by measuring recall and activity rates for apps like Duolingo.

This chapter provides an overview over different classes of spaced repetition algorithms by using a selection of the most significant papers in the field and shows how they perform in terms of recall and activity rate. Further, different conceptual strengths and weaknesses of these algorithms will be presented. For instance, foreign words that sound similar to the native counterpart are easier

[1]https://blog.duolingo.com/global-language-report-2020/

to learn. Thus, one criteria of a spacing repetition algorithm could be whether it can consider the general knowledge base of an individual student.

2 Methodology

To get a general overview over different spaced repetition algorithms, this section provides the search strategy and screening process for finding papers.

2.1 Search strategies

Web of Science and Google Scholar were used to identify current, relevant and often cited paper dealing with spaced repetition optimization. Finding papers for this topic by using key phrases did not lead to a promising result. Thus, we applied another search strategy: the paper "Enhancing human learning via spaced repetition optimization" (Tabibian et al. 2019) was selected and "Connected Papers"[2] was used to identify related papers. Table 1 shows all papers gathered during this process that have something to do with spaced repetition optimization. Two papers also dealing with spaced repetition were excluded because they are only published on ArXiv which means that they are not peer reviewed.

2.2 Screening

Out of these papers, four papers are presented in detail during this chapter. To understand our screening procedure, it makes sense to divide the spaced repetition algorithms into four generations:

1. The Pimsleur Method (1967) is an example of a first generation spaced repetition method. It is the first mainstream use of spacing and lag effects. New items are tested at exponentially increasing intervals. However, the recall is prerecorded on cassete and therefore not adapting to the successful (or unsuccessful) recall of the learner.

2. The Leitner system (see 1). It is still very static but it saves the individual learn success with the help of boxes (see Figure 1).

3. The third generation utilizes AI for optimizing spacing repetition. Since it is used by Duolingo, the most significant paper of this type is Settles & Meeder (2016) The third generation is rule-based but uses machine learning for internal parameter estimation.

[2]https://www.connectedpapers.com/

Table 1: Table of papers dealing with Spaced Repetition Optimization. The number of citations refer to Google Scholar, whereas the number in brackets refer to citations on WebOfScience.

Name	Published	Citations	Year
Enhancing human learning via spaced repetition optimization	Proceedings of the National Academy of Sciences	83 (24)	2019
Deep Reinforcement Learning of Marked Temporal Point Processes	NeurIPS	62	2018
Unbounded Human Learning: Optimal Scheduling for Spaced Repetition	KDD	48	2016
A Trainable Spaced Repetition Model for Language Learning	ACL	153	2016
Improving Students' Long-Term Knowledge Retention Through Personalized Review	Psychological science	202 (63)	2014
DAS3H: Modeling Student Learning and Forgetting for Optimally Scheduling Distributed Practice of Skills	EDM	34	2019
Artificial intelligence to support human instruction	Proceedings of the National Academy of Sciences	12	2019
Large-scale randomized experiments reveals that machine learning-based instruction helps people memorize more effectively	NPJ science of learning	0	2021

4. The fourth generation utilizes non-rule-based AI. The most significant papers of this type (according to citation numbers) are Upadhyay et al. (2018) using Deep Reinforcement Learning and Tabibian et al. (2019) using Stochastic Differential Equation. This generation can generalize across students and items. It can consider the general knowledge base of an individual student.

The fourth paper presented in this chapter is Upadhyay et al. (2021) which conducted a real life, randomized control group study on MEMORIZE, the algorithm behind Tabibian et al. (2019).

3 Results

This section presents the three algorithms presented in the previous section. Because the math and the concept is not too complicated, the algorithm Half Time Regression by Duolingo is presented in more detail than the other algorithms.

3.1 Half Time Regression by Duolingo

The paper of Settles & Meeder (2016) is especially relevant since it is used by Duolingo for spaced repetition. It is a rule-based approach of the third generation. Settles & Meeder (2016) assume exponential decay for the recall rate of an item (Ebbinghaus model of forgetting curve (Ebbinghaus 1885)) and, as the name suggests, they developed a method for regressing the half time h of the recall rate. The method is able to predict a different recall rate for a combination of each student and each item. The probability p that an item is recalled correctly is based on the following formula:

$$p = 2^{-\delta/h} \tag{1}$$

δ denotes the time that has passed since the item was last practiced.

"Let x denote a feature vector that summarizes a student's previous exposure to a particular word, and let the parameter vector Θ contain weights that correspond to each feature variable in x. Under the assumption that half-life should increase exponentially with each repeated exposure (a common practice in spacing and lag effect research), we let h_Θ denote the estimated half-life, given by:

$$\hat{h_\Theta} = 2^{\Theta x} \tag{2}$$

In fact, the Pimsleur and Leitner algorithms can be interpreted as special cases of h_Θ using a few fixed, hand-picked weights." Settles & Meeder (see 2016: page 1851). To get Θ, a loss function trying to minimize the difference between actual probability p and the predicted probability $\hat{p_\Theta}$ is used. With h being the individual half-life of a learners memory. Figure 2 shows the prediction of the forgetting curves versus the actual recall rates (marked by **x**).

The feature set x of Equation 2 can be divided into two categories:

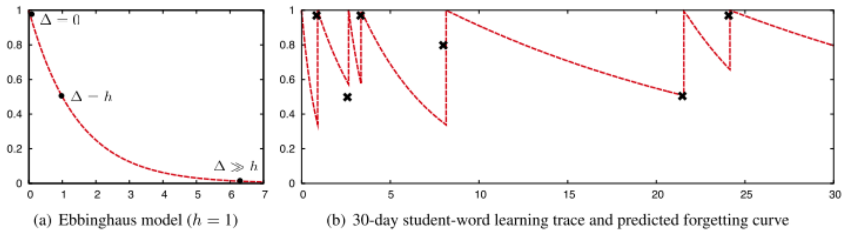
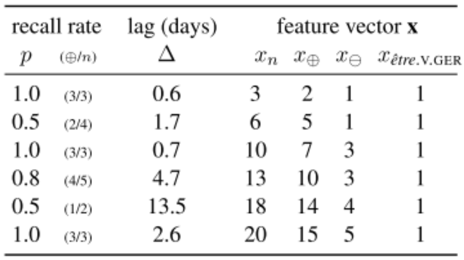


Figure 2: This image is taken from Settles & Meeder (2016) and shows the predicted recall rate of the Ebbinghaus model (a) and the student-word learning trace (b). **x** marks the observed recall rate p for each practice session, and half-life regression aims to fit model predictions $\hat{p}_\Theta$ (dashed lines) to these points.

- individual interaction features: number of times an individual student has successfully and unsuccessfully recalled a flashcard.
- lexeme tag features: capture the inherent difficulty of a flashcard

Figure 3 shows the training instances of an individual student plus the lexeme tag feature $x_{etre.V.GER}$. Each row correspond to a data point of Figure 2.



recall rate		lag (days)	feature vector **x**			
p	($\oplus/n$)	Δ	x_n	$x_\oplus$	$x_\ominus$	$x_{\hat{e}tre\text{.V.GER}}$
1.0	(3/3)	0.6	3	2	1	1
0.5	(2/4)	1.7	6	5	1	1
1.0	(3/3)	0.7	10	7	3	1
0.8	(4/5)	4.7	13	10	3	1
0.5	(1/2)	13.5	18	14	4	1
1.0	(3/3)	2.6	20	15	5	1

Figure 3: This image is taken from Settles & Meeder (2016). Each row correspond to a data point of Figure 2.

The performance of half life regression can be seen in Figure 4

while the results of the controlled user experiments can be seen in Figure 5. The experiments were run in Duolingo's production system involving 1 million users and lasting six weeks. It measured the daily retention rate of users. For evaluation, three retention metrics were used. *Any* measures any kind of user activity (including forum posts etc.), *Lesson* describes whether a user has participated a new lesson and *Practice* for practice sessions. The first row of Figure 5 compares HLR with Leitner system, the second row compares basic HLR with HLR that saves the inherent difficulty of a word.

Model	MAE↓
HLR	**0.128***
HLR -lex	**0.128***
HLR -h	0.350
HLR -lex-h	0.350
Leitner	0.235
Pimsleur	0.445
LR	0.211
LR -lex	0.212
Constant $\bar{p} = 0.859$	0.175

Figure 4: This image is taken from (Settles & Meeder 2016) and describes the Mean Absolute Error (MAE) on historic log data. The model of Settles & Meeder (2016) performs best. Leitner and Pimsleur were used as control groups together with Logistic Regression (LR) and using a constant prediction value of 0.859.

	Daily Retention Activity		
Experiment	Any	Lesson	Practice
I. HLR (v. Leitner)	+0.3	+0.3	*-7.3**
II. HLR -lex (v. HLR)	+12.0*	+1.7*	+9.5*

Figure 5: This image is taken from (Settles & Meeder 2016) and describes the daily retention activity of Duolingo students. Changes are stated in percent.

3.2 Memorize

Current algorithms use simple rule-based heuristics. Memorize, the algorithm used by Tabibian et al. (2019) applies a flexible representation framework called marked temporal point processes. Marked temporal point processes (MTPPs) are a popular framework for modelling asynchronous event data in continuous time like the time in which a vocabulary is learnt. Memorize is a simple, scalable online spaced repetition algorithm. Online in this context means that it can be trained with current live data. It is trained with data from Duolingo and evaluated and proven with synthetic[3] data.

Memorize uses a set of stochastic differential equations with jumps. Stochastical differential equations with jumps are a mathematical model describing the "evolution of [...] random quantities over time"(Platen & Bruti-Liberati 2010).

[3]Procedurally generated

Tabibian et al. (2019) compare the performance of Memorize with two baselines: uniform and threshold-based. The results of the latter are similar to previous results of rule-based algorithms. Only users with at least 30 reviewing events and words reviewed at least 30 times were considered. See Figure 6 for results.

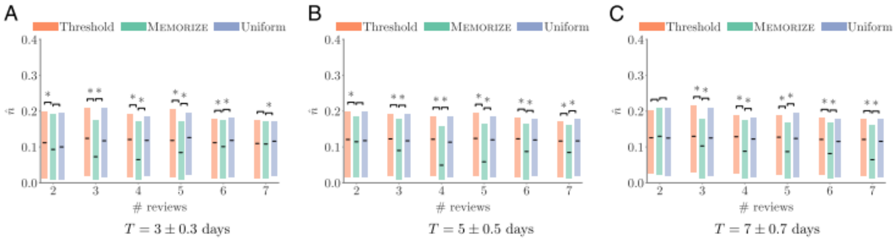


Figure 6: This figure was taken from Tabibian et al. (2019) and shows the empirical forget rate (lower is better). Each triplet of bars in the figures corresponds to the same number of times (# reviews) for approximately the same period of time (T). Boxes indicate 25% and 75% quantiles and crosses indicate median values, where lower values indicate better performance.

Memorize was only evaluated by historic log data. Upadhyay et al. (2021) criticise that MEMORIZE has not conduced a real life, randomized control group study. Therefore, Upadhyay et al. (2021) conducted a large scale ($n = 50,000$) randomized control group study with three test groups themselves: *Select* based on MEMORIZE, *Difficulty* or *Random*. Learners did not know which group they were in. 16.75 million answers to 1900 questions by 50,700 learners in 628,000 study sessions were performed. As can be seen in Figure 7, the select group working with MEMORIZE performed significantly better after 9 days.

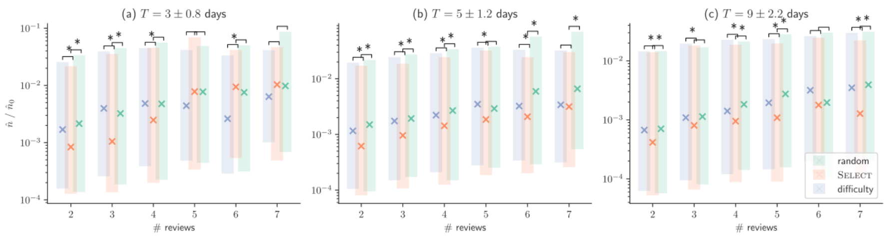


Figure 7: This figure was taken from Upadhyay et al. (2021) and shows the empirical forget rate (Lower is better). Each triplet of bars in the figures corresponds to the same number of times (# reviews) for approximately the same period of time (T). Boxes indicate 25% and 75% quantiles and crosses indicate median values, where lower values indicate better performance.

3.3 Deep Reinforcement Learning for Spaced Repetition

Upadhyay et al. (2018) use Marked temporal point processes (MTPPs) as well. Furthermore, they use Deep Reinforcement Learning for optimization, where both the actions taken by an agent and the feedback it receives from the environment are asynchronous stochastic discrete events. Upadhyay et al. (2018) call their method Temporal Point Process Reinforcement Learning (TPPRL).

Each reinforcement learning setup consist of an environment, an agent and a reward (rewards can also be negative). For the setup of TPPRL applied to spaced repetition[4], the learner acts as the environment, rewards are the scores on a test and the learning platform is the agent deciding when to ask the student to review each item to better prepare them for a test. They use a dataset from Duolingo as training data.

According to the authors, TPPRL outperforms both Memorize and the baseline by large margins (see Figure 8).

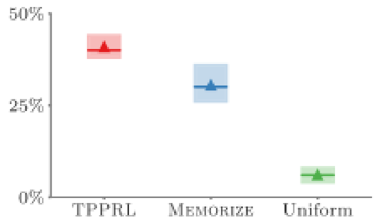


Figure 8: This image is taken from Upadhyay et al. (2018). It shows that the recall rate of their own method TPPRL outperforms the other methods.

However, all Deep Reinforcement Learning model suffer from serious problems: A deep reinforcement learning model is a black box that is very hard to interpret. Using it in production could be problematic, since the cause of occurring problems would not be clear. Furthermore, overfitting the training data is often a problem for deep reinforcement learning models (Zhang et al. 2018).

4 Discussion

The most important evaluation metric is the successful recall rate of an item based upon the number of times the item is reviewed and the duration of the training period. All papers presented in Section 2.2 include this metric. Another

[4]TPPRL can be used in other scenarios than spaced repetition, too

metric used is the daily retention activity for using the learning app. This metric, however, was only used by Settles & Meeder (2016), since there are no reports of the other learning algorithms used in a production setting. Furthermore, it is important to mention that there are two different ways of evaluating a successful recall rate. The obvious one is conducting a randomized control group study. Only Settles & Meeder (2016) and Upadhyay et al. (2021) did so. The other method is measuring the prediction error of the recall rate on historic log data. All papers appeared in conferences or journals. The most prestigious publication of this chapter is Upadhyay et al. (2021) since it appeared in Nature partner journals[5]. However, Settles & Meeder (2016) performed a real life experiment with 1 million real Duolingo users, making the most realistic experimental setting. Besides metrics and scientific standarts, there are other things to consider in real, life production settings: When a new language is introduced to the learning platform and there is no training data. This problem is known as the cold start problem. Only, Settles & Meeder (2016) provide solution for the cold start problem. One version of their algorithm only requires training data from an individual student.

5 Conclusion

Using artificial intelligence for spaced repetition is a good example for the ”silent ubiquitousness” of AI today. Most Duolingo users are probably not aware that AI algorithms are used when they learn their vocabulary. When it comes to selecting an spaced repetition AI algorithm for a productive setting such as DuoLingo, other things than only the recall are very important, too. How the recall rate is evaluated is important as well: the algorithms of Settles & Meeder (2016) and Tabibian et al. (2019) show a good performance in blind study experiments which is important in real life settings. Furthermore, interpretability and solutions for the cold start problem are important as well. This is why Half Life Regression is superior to all other algorithms for a productive setting (which is probably the reason why DuoLingo is using it). However, once there is enough training data, Memorize is interesting as well for a productive setting. This shows that both, a rule-based as well as a ruleless algorithm are interesting for optimizing spaced repetition.

[5] see https://www.nature.com/nature-portfolio/about/npj-series

References

Ebbinghaus, Hermann. 1885. Forgetting curve. *Memory A Contribution to Experimental Psychology. Downloaded from http://encarta. msn. com/media* 461547609. 76157803–1.

Melton, Arthur W. 1970. The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior* 9(5). 596–606.

Platen, Eckhard & Nicola Bruti-Liberati. 2010. Stochastic Differential Equations with Jumps. In *Numerical Solution of Stochastic Differential Equations with Jumps in Finance*, 1–60. Berlin, Heidelberg: Springer Berlin Heidelberg. DOI: 10.1007/978-3-642-13694-8_1. https://doi.org/10.1007/978-3-642-13694-8_1.

Settles, Burr & Brendan Meeder. 2016. A trainable spaced repetition model for language learning. In *Proceedings of the 54th annual meeting of the association for computational linguistics (volume 1: long papers)*, 1848–1858.

Tabibian, Behzad, Utkarsh Upadhyay, Abir De, Ali Zarezade, Bernhard Schölkopf & Manuel Gomez-Rodriguez. 2019. Enhancing human learning via spaced repetition optimization. *Proceedings of the National Academy of Sciences* 116(10). 3988–3993.

Upadhyay, Utkarsh, Abir De & Manuel Gomez-Rodriguez. 2018. Deep reinforcement learning of marked temporal point processes. *arXiv preprint arXiv:1805.09360*.

Upadhyay, Utkarsh, Graham Lancashire, Christoph Moser & Manuel Gomez-Rodriguez. 2021. Large-scale randomized experiments reveals that machine learning-based instruction helps people memorize more effectively. *NPJ science of learning* 6(1). 1–3.

Zhang, Chiyuan, Oriol Vinyals, Remi Munos & Samy Bengio. 2018. A study on overfitting in deep reinforcement learning. *arXiv preprint arXiv:1804.06893*.

Chapter 8

Personalized learning experiences

Shantanu Audichya & Elena Korovina

Personalized learning has emerged as a solution to the intersection of information overload and time spent on search by learners. The goal of personalized learning is to improve the learning effectiveness and efficacy of learners. This is achieved by allowing the student to follow the desired learning material at their own pace, which in turn enhances their motivation for learning. The objective of this review is to analyze the research on AI-based personalized learning in the field of education. We covered algorithms from across the spectrum, from traditional machine learning algorithms such as content based filtering and collaborative filtering at one end, to deep learning models such as neural networks at the other. While analyzing, we experienced the change of landscape, in the framework of customised learning, over a period of advancement in the field of artificial intelligence. We realised that the personalized experience is no longer restricted on the basis of similarity of learning contents. With all the improvement of various artificial intelligence algorithms, induced in the field of EdTech. We came across literature, where authors leverage neural network architecture to capture implicit features of learners as well.

Keywords: Recommendation System, Artificial Intelligence, Personalized Learning, Education, e-learning

1 Introduction

The advancement of technologies and easy access to the internet resulted in the explosion of online content. This overload of information, in turn, led to the democratisation of education. However, the information explosion also resulted in the problem of relevancy of online materials. As, to access the information, a learner is only required to have a device with Internet connection. The vastness of online content is a challenge especially in the domain of education.

In this chapter, we are analysing Artificial Intelligence based systems for personalized learning that aim to provide students with adaptive recommendations. Thus, they create individual learning paths and aim to fulfill the requirements of different learners. These AI-based solutions can help in shifting the paradigm from one-size-fits-all approach to personalized online courses, which would consider learners abilities and requirements. In personalized education, all students are believed to have equal capacity to learn if they are provided with appropriate tools.

Our scope of research is restricted to the study of concepts related to various advancements in the domain of education. Specifically, we are trying to answer from the perspective of the following open ended questions:

- What are the advances the Educational Technology (from here on out referred to as EdTech) industry has seen from the customized content point of view?

- What does the future hold in amalgamation of e-learning space and artificial intelligence?

2 Methodology

We initiated our research with the digital libraries- ACM and IEEE. In order to supplement our research, we also considered the citation database- Scopus. Since our intention was to carry out holistic research, we referred to the web search engine Google Scholar as well. Firstly, the initial keywords such as ”Recommender system”, ”e-learning”, ”education” and ”personalized learning” were used as a part of preliminary search and to understand the common keywords present in the papers related to the topic of our interest. On the basis of the output of preliminary search, the search process was updated by adjusting the ”Advanced Search” option. With the usage of Advanced Search option, the flexibility of using logical operators with the keywords was leveraged.

As part of the updated search strategy, the keywords were enhanced to filter out more relevant results. The updated keywords used were: ”recommender* system” OR ”personalized learning” OR ”collaborative filtering” AND ”education” OR ”e-learning”. In terms of the publication time range, our research was restricted between the years 2012 and 2022. The decision of selecting this time period was based on the two main assumptions: (1) The year 2012 marks the emergence of the concept ”Massive Open Online Course” (MOOC’s). It is the concept aimed at unlimited participation and open access via the web. So, the assumption was made that the online content became prevalent around the year 2012. (2) According to Chong et al. (2020), the introduction of big data and student profiling with learning analytics became a dominant focus in 2010-2019.

Table:1 specifies the number of records from mentioned databases and the search term used. These numbers describe the initial output from each databases. The records numbers were further filtered during our second stage of scrutiny.

Table 1: Initial search strings

Database	Search terms	Records
IEEE	(”recommender system” OR ”personalized learning” OR ”collaborative filtering” OR ”content based filtering”) AND (”education” OR ”formal education” OR ”e-learning” OR ”students”)	66
ACM	(”recommender system” OR ”collaborative filtering” OR ”personalized learning” OR ”content based filtering”) (”education” OR ”e-learning” OR ”formal education” OR ”student” OR ”edtech”)	36
Scopus	”recommendation system” OR ”collaborative filtering” OR ”content based filtering” OR ”personalized learning”) AND (”EdTech” OR ”education” OR ””e-learning” OR ”formal education”)	439
Google Scholar	”personalized learning” OR ”artificial intelligence” OR ”high school” OR ”education”	31
Total Records		572

At the beginning, a total of 572 records were selected by implementing the search strategy on above mentioned main databases, namely - IEEE, ACM, Scopus and Google Scholar. Once the initial records were obtained, the first step of filtering was implemented. In the first step of filtering, all duplicate records present in more than one databases were removed. As a part of second filtering criteria, records were selected on the basis of domain of our interest i.e. personalised learning in format education for students. Only the records in English were selected for further scrutiny. As a result, this paper is based on 9 scientific articles.

3 Results

We structured our results on the basis of artificial intelligence techniques leveraged by different papers. We initiated the process with the most traditional machine learning techniques by capturing the different views on this topic. After that, we moved through the more advanced techniques like deep neural networks.

3.1 Traditional Machine Learning

The inclusion of various machine learning based algorithms in the education domain has changed the landscape of the EdTech industry. In the write up on personalization in education (S.S. et al. (2021)), the authors argue for the variation of recommendation systems such as content-based, collaborative filtering and hybrid recommendation systems. In order to explore the implicit factors such as social media interactions of learners, Dwivedi & Bharadwaj (2012) explored the association among retrieval resources, based on the learners social media interactions. Figure 1 illustrates the information structure for resource recommendation in online social networks. This architecture enhances the recommendation coverage and handles the sparsity problem (Problem refers to the phenomenon of not observing enough data in the dataset).

As to further make the personalized system holistic, the authors of the paper K. & N. (2019) proposed a two-phases model - ”preprocessing” and ”prediction”, for personalised learning on the basis of URL search. In the first phase, the process begins by ”preprocessing” the URL to extract the keywords in the form of tokens. For the second phase, according to the keywords extracted, the new ”prediction” is made for the suggested URL. For this, the authors used Random Forest Algorithm.

Other algorithms were also explored in the paper Lin et al. (2013). Here data mining (an application of decision trees) emerged as the best way to provide an adaptive web-based environment for learning creativity. In their personalize creativity learning system (PCLS), they integrated personalized and game-based learning approaches. Creativity was trained by prompting users to give original solutions for stated problems in a game scenario. Decision trees evaluated student’s performance and optimized individual learning experience. Statistical algorithms were deployed to keep track of information about an individuals learning process. This insight into the cognitive capability of the students can be presented to the teacher for a future analysis.

PCLS takes personal needs of the student into account and examines their individual learning style. Thus, the individually most effective learning path can be recommended by the system.

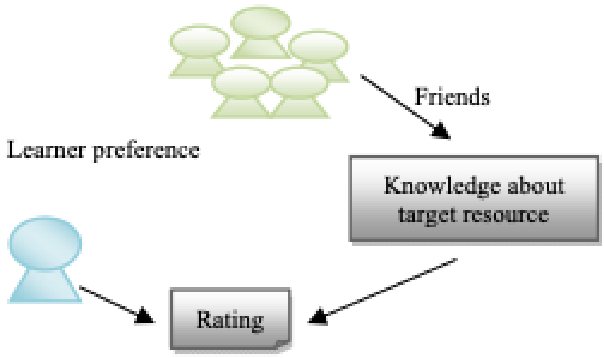


Figure 1: Information structure for resource recommendation (Dwivedi & Bharadwaj 2012)

3.2 Deep Learning

The latest innovations in machine learning have led to various deep learning techniques. These techniques are a powerful tool for performing all sorts of tasks. That is why research in the educational field has also tried to incorporate deep learning into systems for personalized education.

McCarthy et al. (2020) proposed the automated tutor iSTART (and its later modifications iSTART-ME and iSTART-3) which aims to enhance reading comprehension of the learners by the means of self-explanation. Self-explanation is the ability to explain the reading material to oneself by paraphrasing and building connections to one prior knowledge as well as other parts of the text. iSTART provides students with practice for different comprehension strategies in a personalized manner. This technology favors learner's unique point of view and keeps track of their personal learning progress.

iSTART-ME was modified by a gaming environment with a customizable player avatar. This further increased user's engagement levels.

Moreover, iSTART-3 aimed to propose texts to the individual's skills and abilities. Adaptive text selection motivates students and results in a better performance and learning outcomes.

This tutoring system utilizes natural language processing (NLP) technologies in order to assess students self-explanations and give an individualized feedback to them. Adaptive logic techniques were implemented in order to tailor text difficulty to the learners based on their previous performance.

The paper by Chanaa & El Faddouli (2018) proposes an automated processing model with ability to solve learning problems without requiring human interventions. The authors suggest a personalised system that considers variables related to the learners themselves, such as learning speed, duration of study etc. instead of variables related to the content. The model architecture can be seen in the figure below:

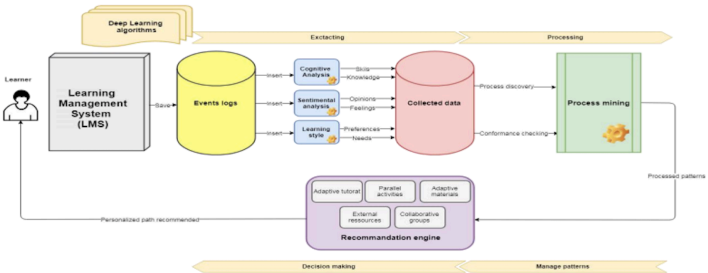


Figure 2: Proposed personalised model architecture (**Chanaa, Faddouli 2018**)

Nowadays, many online education platforms utilize traditional data mining methods such as Collaborative Filtering (CF). Despite the development and contributions of many recommender systems, diverse deep learning models for personalized recommendations are being explored because of problems such as sparsity and scalability. Sparsity refers to the phenomenon of not observing enough data in the dataset and scalability is the ability to incorporate data complexities. For example, a study by Q. & Kim (2021) proposes a novel deep learning based recommender system (DECOR). The model captures high-level user behaviors and course attribute features. The authors suggest that it could potentially reduce information overload, solve the high-dimensional data sparsity problem, and achieve high feature information extraction performance. The model architecture is showcased in Figure 3.

The overall architecture has 3 main modules:

- **User Behavior Extractor Module (UBE)** - This module captures high-level user behavior feature and learns complex relationships.
- **Course Attribute Extractor Module (CAE)** - This module learns high-level course attribute features.

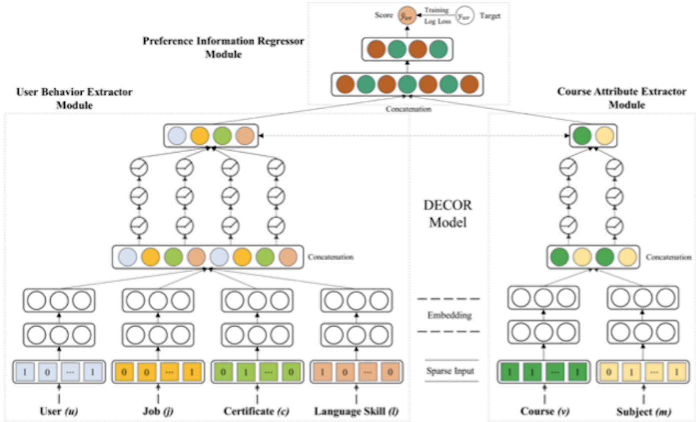


Figure 3: DECOR Model Framework (**Li**, **Kim** **2021**)

- **Preference Information Regressor Module (PIR)** - This module is responsible for integrating high-level representation features obtained from UBE and CAE into an end-to-end process to implement ensemble learning.

As a more targeted platform, authors Rad et al. (2018) created Cloud-eLab, which is an open and interactive cloud-based learning platform. Its aim is to encourage AI Thinking, learning of data analytics and network security. The authors describe Cloud-eLab as a tool for learning how to construct scientific experiences. This is achieved by allowing students to have control over the learning process, being able to go through the material at their own pace. Besides, each user has access to personalized content. This platform utilizes various techniques to encode data and to evaluate answers given by the user. Additionally, advanced machine learning techniques are utilized, allowing for emotion recognition, face detection, audio and text analysis.

4 Conclusion

The incorporation of machine learning based algorithms, especially in the domain of education, has garnered support from the wider audience. From our analysis, we have observed that the use of Recommendation Systems forms the basis of personalized learning. The performance of recommendation algorithms can be further enhanced. Examples include but are not limited to machine learning techniques (such as Random Forest Algorithm, decision trees) and deep learning.

While traditional learning systems neglect optimization of the learning materials based on student needs, personalized learning approaches provide different learning strategies. It takes into account the fact that students come from divergent knowledge backgrounds, as well as have different learning abilities, skills and fields of interest. Thus, customization of the learning path leads to enhanced motivation and learning efficacy.

However, we have also observed the limitations of personalized learning. For instance, in countries where the user data laws are stringent, less studies focusing on Web-based personalized learning systems were found.

We would like to conclude by pointing out that the amalgamation of machine learning techniques applied in the educational field presents a wide scope of possibilities to experiment and has a great potential of enhancing the learning experience.

References

Chanaa, Abdessamad & Nour-eddine El Faddouli. 2018. Deep learning for a smart e-learning system. In *Proceedings of the 2nd international conference on smart digital environment*, 197–202. New York, NY, USA: Association for Computing Machinery. DOI: 10.1145/3289100.3289132. https://doi.org/10.1145/3289100.3289132.

Chong, Guan, Mou Jian & Jiang Zhiying. 2020. Artificial intelligence innovation in education: a twenty-year data-driven historical analysis. *International Journal of Innovation Studies* 4. 134–147. DOI: 10.1016/j.ijis.2020.09.001.

Dwivedi, Pragya & Kamal K. Bharadwaj. 2012. E-learning recommender system for learners in online social networks through association retrieval. In *Proceedings of the cube international information technology conference*, 676–681. New York, NY, USA: Association for Computing Machinery. DOI: 10.1145/2381716.2381846. https://doi.org/10.1145/2381716.2381846.

K., Chaudhary & Gupta N. 2019. E-learning recommender system for learners: a machine learning based approach. *International Journal of Mathematical, Engineering and Management Sciences* 4. 957–967. DOI: 10.33889/IJMEMS.2019.4.4-076.

Lin, Chun Fu, Yu-chu Yeh, Yu Hsin Hung & Ray I Chang. 2013. Data mining for providing a personalized learning path in creativity: an application of decision trees. *Computers & Education* 68. 199–210. DOI: 10.1016/j.compedu.2013.05.009.

McCarthy, Kathryn S., Micah Watanabe, Jianmin Dai & Danielle S. McNamara. 2020. Personalized learning in istart: past modifications and future design. *Journal of Research on Technology in Education* 52:3. 301–321. DOI: 10.1080/15391523.2020.1716201.

Q., Li & J. Kim. 2021. A deep learning-based course recommender system for sustainable development in education. *International Journal of Mathematical, Engineering and Management Sciences*. DOI: 10.3390/app11198993.

Rad, Paul, Mehdi Roopaei, Nicole Beebe, Mehdi Shadaram & Yoris A. Au. 2018. Ai thinking for cloud education platform with personalized learning. In *Proceedings of the 51st hawaii international conference on system sciences*, 879–888. Hilton Waikoloa Village, Hawaii. DOI: 10.24251/HICSS.2018.003.

S.S., Kundu, Sarkar D., Jana P. & Kole D. 2021. *Personalization in education using recommendation system: an overview*. Cham: Springer Science and Business Media Deutschland GmbH.

Chapter 9

Data sets in AI in education

Erika Monserrat Angelescu & Jacqueline Näther

The data necessary for the development of AI is especially protected in Education. This complicates the development of AIs in Education (AIED). To counteract this, many developers are already researching improved anonymized data sets for the development of AIED. Their handling with topics, such as purely anonymous data with upgraded AIs, modified data sets using data augmentation and the creation of purely synthetic data sets. We will analyze these approaches regarding strengths and weaknesses. In addition, further approaches which have also improving ideas to use anonymized data sets or improving the results of predicting students performance are presented. In the end, combining different approaches can improve the handling of AIED while also protecting students privacy. We come across a problem between ideal data and ideal models in which we conclude that more studies are to be performed to get a better understanding of this problems. **Keywords:** Data Availability, Data Sets, Privacy Protection, Artificial Intelligence

1 Introduction

Most of today's applications that make use of artificial intelligence require a great amount of data. Therefore, the availability and quality of large data sets is important and necessary to discuss. Access to this data has become easier with the digital change in the field of education. But with the collection and storage of data come the difficulties of protecting individuals' personal information. Various bodies drafted laws for the protection of privacy and data protection. State institutions must especially deal with vulnerable people, such as children. As an example, the United States' regulations for recording student education are governed by the *Family Educational Rights and Privacy Act (FERPA)* (U.S. Department of Education 2021) to ensure privacy for students.

With the existing regulations, the students data had to be anonymized even more. This means that parts of the existing data, such as name and age, have to be removed. This not only reduces the data set itself, but also the number of possible features that can be formed from it. The smaller the data set and the fewer features it provides, the more difficult it is to train an AI with it and achieve high levels of accuracy. With small data sets with little diversity, there is for example often a risk of overfitting. In order to counteract the problems mentioned, various approaches are already being investigated. These approaches are explained as part of our investigation. The approaches range from simple anonymization and upgrading the AIs, removing clearly assignable information via modified data (data augmentation) to the development of purely synthetically produced data. In particular, strengths and weaknesses of these approaches are analyzed and possible improvements are further discussed.

We will not discuss in detail what type of data sets are used and what kind of information or features are build. We concentrate on how data sets were modified for each experiment, how the authors worked with them and rather analyze if the data sets could have brought a problem to the results shown in the experiments. We will try to expose what could have made the data set better for a more accurate results.

We first describe the methodology of how the discussed papers were found. We then go into the explanation and present the results of our investigations. Finally, other approaches are considered, which also have an influence on the data sets used in AIED. In our conclusion we combine the different ideas of the mentioned approaches and give an idea of what combined approaches could probably achieve.

2 Methodology

For reasons of cost and the scarce availability of software tools, we decided to use a manual study search strategy. At first, we looked for information material on the topic of AIED that is available online and offline in the library of the Osnabrück University. After finding some seminar books and papers on the topic of AIED, we defined further criteria for the papers and their sources.

- The papers need to be in English.
- Papers should be recent, the more recent, the greater impact it will have to our paper.

- The papers should be examined and published via an official committee or conference.

The publications from the *International Conference on Artificial Intelligence in Education*[1] suggested as a good source of papers for our proposed topic, since these papers were published by a conference committee and thus checked for validity and can be considered trustworthy. Therefore, we base our investigations on these papers. However, we also decided to include suitable parts from papers published elsewhere in order to provide a more comprehensive overview in our analysis.

From the *International Conference on Artificial Intelligence in Education* we first came across the book from the conference held in 2015 (Artificial Intelligence in Education, 17th International Conference, AIED 2015, Madrid, Spain, June 22-26, 2015. Proceedings). We then thought of searching for a more recent conference from the same topic and came across the conference held in 2021 which has two volumes (Artificial Intelligence in Education, 22nd International Conference, AIED 2021, Utrecht, The Netherlands, June 14–18, 2021, Proceedings, Part I), and (Artificial Intelligence in Education, 22nd International Conference, AIED 2021, Utrecht, The Netherlands, June 14–18, 2021, Proceedings, Part II). After deciding on these two conference books we started to select specific papers out of the books. To do this, we first looked at all the titles in the table of contents of the publications. At this point, we marked all those whose title can indicate work with interesting content on the subject of data sets, processing, development and investigation. Important keywords were:

- Data, Data Analysis, Data System, Data Augmentation or Data Set
- Development
- Difference between Participants
- Feature
- Prediction
- Privacy
- Protection
- System

[1]You can find out more about the conference on it’s homepage: https://iaied.org/

From 137 papers of the Conference held in 2015 we marked 16 as interesting papers and from the 125 papers of the Conference held in 2021 we marked 24 papers as interesting. These 40 papers were considered for further investigation.

The further investigation consisted of reading the abstracts of the 40 selected papers and, based on the content, we confirmed whether the paper can provide relevant information for our study. Afterwards we identified 15 papers that offer potential for our study regarding data in AIED. We carefully read these ones and could eliminated other eight that were not satisfactory to the topic we were focusing on. This led us to seven interesting articles on which further research is based. After, we decided to choose three papers to discuss in detail and further invest into individual parts of the other five papers. All steps were carried out with interim results and the final number of selected papers is shown graphically in Figure 1.

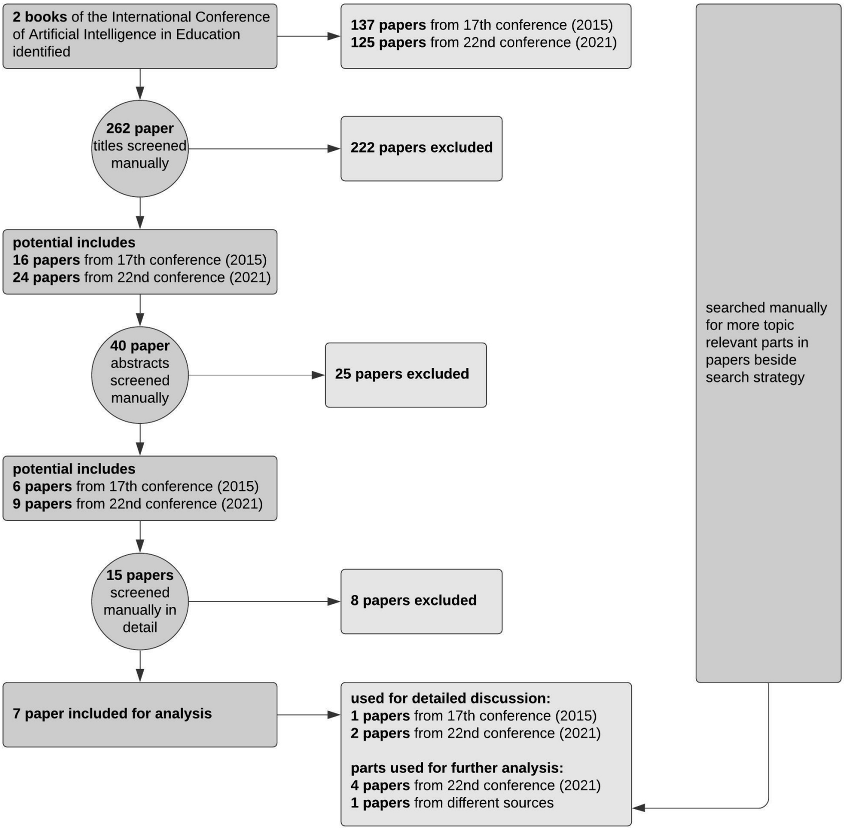


Figure 1: Graphical representation of the search strategy.

3 Results

In the next section, we will first introduce the studies and explain how they use the data sets. The results of the studies from the papers will also be described. Then, we discuss the strengths and weaknesses of the data sets. Finally, through a brief overview of other works, we will present an outlook on possible improvements and suggestions for future developments of reliable data sets.

3.1 General approaches

This section we concentrate on presenting three studies and their results in base of their data set. The first study will describe some more about the student performance prediction task. The other ones will discuss the same task, but concentrate more on the data set and the modifications of them.

In the first study (Koprinska et al. 2015), the researchers wanted to predict students' success in the midterm exam (short term predictions) and in the final exam (long term predictions). Koprinska et al. (2015) took a total of 224 students from a first year programming course. They created three groups regarding success in the evaluation: High-level students with an exam mark of [75, 100], Average-level students with an exam mark of [50, 74] and Low-level students with a failed exam. Results showed for short term predictions a 66.52% of accuracy, meaning that two students out of three were correctly classified into a group of success, but also stating that one out of three students would be wrongly classified; and for long term predictions a 72.69% having roughly three students correctly classified out of four. First of all, this study shows that a prediction of student performance can be carried out with the help of an AI.

A more recent paper also deals precisely with this objective, but uses a different data set for this purpose. As in the previous study, six years later Bell et al. (2021) also meant to predict student performance. The data set used in this study was anonymous data from 50 students. In the previous study (Koprinska et al. 2015) a less good prediction result can already be determined. Since the data set used in this study is about five times as small, one might assume that even worse prediction results would have to be expected here. Moreover, the anonymization of the data made it even more difficult to accomplish this study since the features had to be reduced in order to maintain the information anonymous. But Bell et al. (2021) carried out a new approach to use the data set at their maximum. Their approach was to artificially expand the data in order to be able to achieve a higher accuracy in the prediction. Here, the existing data set is expanded with new features. On the one hand, the augmented data was derived or calculated from the student data, and on the other hand, it was derived and added from further student-independent information, like class averages or general percentage of french students repeating high-school years (Bell et al. 2021). Finally, three categories of augmented data were created. As the initial data set, data from the 50 students to predict the performance of the students. They have also tested the accuracy of their classification with other combinations of the augmented data. For this purpose, the initial data set was combined with one of the augmented data sets, with two of the augmented data sets and with all three augmented data sets. The results of this study turned out as inconclusive for only 50 students using only the initial data set. Nevertheless, combining the given data set with only one of the augmented data sets did not lead to any improvement in accuracy. Increased accuracy can be seen with the combinations of the initial data and metrics and counters but it makes a more notable increase in accuracy if all four data sets are combined with each other. Since both approaches, Koprinska et al. (2015) and Bell et al. (2021), test their data set with the help of different algorithms, a direct comparison is difficult. Nevertheless, it turns out that even with only about a quarter of students and the modification of the data set, the accuracy of the predictions is at least as good and in some cases even better than in Koprinska et al. (2015).

In the same year, Bautista & Inventado (2021) expand the idea of using artificially generated data to train AIED. They create a generative adversarial network (GAN) that produces synthetic data out of given real data. The newly generated data set can not be tracked back to the actual given data of real students. This presents a completely new development that could protect students privacy while still receiving a reliable data set for the development of AIED. In the first step of their analysis, the error residuals between the data sets are calculated. On average, only minor differences can be found here. This is also considered necessary, as this reflects that there will be no overfitting and thus it is ensured that the GAN has not created any samples from the original data set. However, some issues were encountered after the number of features in the data set was significantly increased. The GAN becomes unstable and the previously examined value for the average residual also increases. The results of the synthetic data set are not as precise as the original data set anymore. Therefore, this GAN has to be further developed before it can produce synthetic data sets for data including lots of features that can be used as equivalent to real student data.

In comparison, all studies have tried to generate a data set that is sufficiently meaningful so that a prediction is possible. In the last-mentioned study (Bautista & Inventado 2021), the use of the data set was not checked. This is still a missing measure to be able to fully validate the results. Overall, it has been shown that artificially produced data sets can be a good alternative to real data. This is a sustainable approach that should be further explored in the future.

3.2 Advanced approaches

In addition to previously described approaches, there are other methods which can expand the ideas of the previous studies. They do not concentrate on the same task anymore, but discuss approaches to work with data sets that are concerned with securing privacy. The approaches described below could be combined in future studies with the results from the field of artificial data development in Education. Theoretically possible symbioses that arise from this will be explained below. However, these must also be empirically investigated in the future.

So far in the studies, it has been assumed that a large number of features in a data set including a lot of data from different students can lead to good results. It is typical for these studies, however, that marginal groups are viewed as outliers, as these rarely appear in the data sets. It is possible that they are no longer even available in the case of purely synthetically produced data sets, like in Bautista & Inventado (2021). However, so that these fringe groups are also taken into account in the future, Yun et al. (2021) developed the Self-paced Graph Memory Network (SPGMN) where they merge Self-paced Learning (SPL) (Meng et al. 2017) with Graph Memory Networks (GMN) (Hosein Khasahmadi et al. 2020) to get more robustness to the GMN. As a result, the accuracy of the prediction is improved. In this study, the authors use a large data set with data from 600 students. In order to be able to use the data in the SPGMN, it is created in the form of graphs so in that way all values of a student correspond to a graph. The results show that the accuracy of the prediction of the SPGMN is significantly better than that of the usual network variants. In addition to the Student Learning Performance Prediction (SLPP) task, the Abnormal Students Detection (ASD) method was then used to check whether and to what extent abnormal students were detected. Theoretically, it is now conceivable to carry out a study in which the data set used is partially or fully augmented, as described in Bell et al. (2021) and Bautista & Inventado (2021). It is then particularly important to check, whether comparable results could also be achieved with the ASD method with the synthetic data sets. The partially or fully synthetic data set can then be checked in an additional way. It plays a particularly important role to check, how realistic it is to create such a method and what percentage of the correct predictions would be obtained in the end.

Ouyang et al. (2021) propose a Position-aware Self-Attentive Knowledge Tracing (PAKT) model so that when having a chronological sequence of exercises and its results, one can predict the probability of answering a new exercise correctly. They develop PAKT with the help of Self-Attentive Knowledge Tracing (SAKT) (Zhang et al. 2017) and Deep Knowledge Tracing (DKT) (Piech et al. 2015). To compare the results of former methods and their own, Ouyang et al. (2021) ran some tests with ASSIST2009[2], Simulated-5 (Piech et al. 2015) and ASSIST2015[3] data sets. They came to the conclusion that PAKT was the most effective between all Models taken in consideration. For further reading in their development see Ouyang et al. (2021).

Čechák & Pelánek (2021) introduce a new problem that also has to be addressed. They develop an ideal model to show students skill distribution and its affects over estimates and compare multiple models with it to detect the most accurate model. Therefore they create a model based on variations of the commonly used model the Additive Factor Model (AFM) (G. Durand et al. 2017). For more information to their own model we recommend a more deep read of their paper since we will not get into many details on how they created their model. We just concentrate on their results. They came to the conclusion that although the model was ideal, the original AFM was more accurate in results than their ideal one.

3.3 Other approaches

After describing some studies specialized in how to create different kind of data sets, and how to improve these data sets to become more reliable, some alternatives will be discussed. In this section we present Baker et al. (2021) and their five main ideas to what should be improved for a more reliable model. These ideas are presented in Table 1 and potential steps are also shown.

Baker et al. (2021) introduce a connection between two learning systems that not only makes a more accurate prediction of students' future achievements but could also help students learn in a more efficient way. They divided this process into five challenges and introduced ways of how to solve this challenges like shown in Table 1 These parameters could be useful to include in future studies.

Another important topic should be considered and discussed, else this could become an issue in the future. This would be privacy protection and the morality of using data from actual students for research. In another study (Čechák &

[2] See for more information: https://sites.google.com/site/assistmentsdata/home/assistment-2009-2010-data/skill-builder-data-2009-2010

[3] See for more information: https://sites.google.com/site/assistmentsdata/home/2015-assistments-skill-builder-data

Table 1: Challenges to successfully connecting two learning systems

Challenge	Description	Potential Steps
Connection	Has to be logical and digital	*System-weighted averaging*: Take average of system's estimate, own system estimate weights more than the other
Mapping Related Constructs	Student models have to be similar or related	*System and evidence quantity weighted averaging*: take average of system's estimates, system weights other systems evidence by amount of evidence with penalization of not local system
Evidence Integration	Possibility to integrate data from the other system	*Performance Factors Analysis (PFA)* (for more information see Pavlik Jr et al. (2009))
A Good Reason	Practical reason for connecting student models	*Bayesian Network*
Demonstration of Benefit	Has to make a difference to in students' behavior	*Deep Knowledge Tracing +*(for more information see Yeung & Yeung (2018))

Pelánek 2021) they tried to use simulated data instead of data from actual students but received worse predictions with their perfect model and simulated better data than with, how they described it, data without considering implementation details. If this is a problem of the model or of the simulated data is unclear. A similar study but instead of using simulated data benefit from actual data from students could be a possible solution to solve the unclear statement.

4 Conclusion

Various studies were presented that examined the challenges of using anonymized data sets to predict student performance with AI. All approaches from purely synthetically produced data to modified data sets to simple data sets with many features prove to be useful.

To conclude, creating the ideal data set or the ideal model may not be the best solution for more precise results, as it can be seen in Čechák & Pelánek (2021). Nevertheless, further research in this domain should be made to understand what part of the ideal model generated the problem to worse predictions. Two possibilities are created in this scenario, (1) the possibility of simulated data having influenced in the results of Čechák & Pelánek (2021) and therefore needing data from real students to make reliable research; or in the case that simulated data and real data would react in the same way when repeating the study and therefore (2) there is no such thing as an ideal model we can work with. The latter case, would solve some issues like the privacy of students, in the sense that no more data from students need to be used to make artificial intelligence work. However, if it is indeed the case of an ideal model not working as one expected to, there could be a possibility that the ideal data would also not work as ideally as it was predicted to. Further studies in this topic should be made to come up with a better solution.

The different approaches discussed were able to provide solutions that can combat the problems of small data sets with additional few features. In order to be able to make a qualified statement about the applicability of the solutions presented and possible combinations of the different solutions, further studies must be carried out. With these, the occurrence of overfitting in particular must be examined while at the same time ensuring high accuracy for students performance prediction tasks.

In summary, it can be said that there are many different approaches to countering the problem of using an anonymous data set. In our investigation, only small insights into a limited number of studies could be given. This should be further deepened and expanded in the future. Ideas and suggestions for further improvement were identified and presented. In further investigations, these should be tested and used profitably in the field of AI in Education.

References

Baker, Ryan S., Bruce M. McLaren, Stephen Hutt, J. Elizabeth Richey, Elizabeth Rowe, Ma. Victoria Almeda, Michael Mogessie & Juliana M. AL. Andres. 2021. Towards sharing student models across learning systems. *Artificial Intelligence in Education, 22nd International Conference, AIED 2021, Utrecht, The Netherlands, June 14–18, 2021, Proceedings, Part II* 22. 60–65. DOI: 10.1007/978-3-030-78270-2_10.

Bautista, Peter & Paul Salvador Inventado. 2021. Protecting student privacy with synthetic data from generative adversarial networks. *Artificial Intelligence in Education, 22nd International Conference, AIED 2021, Utrecht, The Netherlands, June 14–18, 2021, Proceedings, Part II* 22. 66–70. DOI: 10.1007/978-3-030-78270-2_11.

Bell, Timothy H., Christel Dartigues-Pallez, Florent Jaillet & Christophe Genolini. 2021. Data augmentation for enlarging student feature space and improving random forest success prediction. *Artificial Intelligence in Education, 22nd International Conference, AIED 2021, Utrecht, The Netherlands, June 14–18, 2021, Proceedings, Part II* 22. 82–87. DOI: 10.1007/978-3-030-78270-2_14.

Čechák, Jaroslav & Radek Pelánek. 2021. Better model, worse predictions: the dangers in student model comparisons. *Artificial Intelligence in Education, 22nd International Conference, AIED 2021, Utrecht, The Netherlands, June 14–18, 2021, Proceedings, Part I* 22. 500–511. DOI: 10.1007/978-3-030-78292-4_40.

Conati, Cristina, Neil Heffernan, Antonija Mitrovic & M. Felisa Verdejo. 2015. *Artificial intelligence in education, 17th international conference, aied 2015, madrid, spain, june 22-26, 2015. proceedings* (Lecture Notes in Computer Science). springer.

G. Durand, C. Goutte, N. Belacel, Y. Bouslimani & S. Leger. 2017. Review, computation and application of the additive factor model (afm). *National Research Council Canada (2017).* 1–23. DOI: 10.4224/23002483.

Hosein Khasahmadi, Amir, Kaveh Hassani, Parsa Moradi, Leo Lee & Quaid Morris. 2020. Memory-based graph networks. DOI: arXiv:2002.09518.

Koprinska, Irena, Joshua Stretton & Kalina Yacef. 2015. Predicting student performance from multiple data sources. *Artificial Intelligence in Education, 17th International Conference, AIED 2015, Madrid, Spain, June 22-26* 17. 678–681. DOI: 10.1007/978-3-319-19773-9_90.

Meng, Deyu, Qian Zhao & Lu Jiang. 2017. A theoretical understanding of self-paced learning. *Information Sciences* 414. 319–328. DOI: 10.1016/j.ins.2017.05.043.

Ouyang, Yuanxin, Yucong Zhou, Hongbo Zhang, Wenge Rong & Zhang Xiong. 2021. Pakt: a position-aware self-attentive approach for knowledge tracing. *Artificial Intelligence in Education, 22nd International Conference, AIED 2021, Utrecht, The Netherlands, June 14–18, 2021, Proceedings, Part II* 22. 285–289. DOI: 10.1007/978-3-030-78270-2_51.

Pavlik Jr, Phil, Hao Cen & Kenneth Koedinger. 2009. Performance factors analysis - a new alternative to knowledge tracing. *Frontiers in Artificial Intelligence and Applications* 200. 531–538. DOI: 10.3233/978-1-60750-028-5-531.

Piech, C., J. Bassen, J. Huang, S. Ganguli, M. Sahami, L. J. Guibas & J. Sohl-Dickstein. 2015. Deep knowledge tracing. *In Proceedings of the 28th International Conference on Neural Information Processing Systems* 1. 505–513. DOI: 10.5555/2969239.2969296.

Roll, Ido, Danielle McNamara, Sergey Sosnovsky, Rose Luckin & Vania Dimitrova. 2021a. *Artificial intelligence in education, 22nd international conference, aied 2021, utrecht, the netherlands, june 14–18, 2021, proceedings, part i* (Lecture Notes in Computer Science). springer.

Roll, Ido, Danielle McNamara, Sergey Sosnovsky, Rose Luckin & Vania Dimitrova. 2021b. *Artificial intelligence in education, 22nd international conference, aied 2021, utrecht, the netherlands, june 14–18, 2021, proceedings, part ii* (Lecture Notes in Computer Science). springer.

U.S. Department of Education. 2021. *Family educational rights and privacy act (ferpa)*. https://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html.

Yeung, Chun-Kit & Dit-Yan Yeung. 2018. Addressing two problems in deep knowledge tracing via prediction-consistent regularization. DOI: 10.1145/3231644.3231647. https://doi.org/10.1145/3231644.3231647.

Yun, Yue, Huan Dai, Ruoqi Cao, Yupei Zhang & Xuequn Shang. 2021. Self-paced graph memory network for student gpa prediction and abnormal student detection. *Artificial Intelligence in Education, 22nd International Conference, AIED 2021, Utrecht, The Netherlands, June 14–18, 2021, Proceedings, Part II* 22. 417–421. DOI: 10.1007/978-3-030-78270-2_74.

Zhang, J., X. Shi, I. King & D.Y. Yeung. 2017. Dynamic key-value memory networks for knowledge tracing. *Proceedings of the 26th International Conference on World Wide Web, WWW 2017* 22. 765–774. DOI: 10.1145/3038912.3052580.

Chapter 10

Automatic lecture transcription

Nele Daske, Raia Abu Ahmad & Febryeric Malsom Parantean

Transcribing lectures automatically has been a research area within the Artificial Intelligence in Education field since the late 1990s. However, not a lot of research has been dedicated to this field since. In this paper, we conduct a systematic review of 11 papers that deal with this field which we believe can be very beneficial, especially in light of the recent pandemic. We describe the methods we used in conducting our review, as well as the overall language models, acoustic models, and datasets used for developing automatic lecture transcription systems. We conclude that the field is very under-researched, and more novel approaches and techniques should be applied to improve the quality of transcripts, especially considering the advances in natural language processing applications in recent years.

Keywords: lecture transcription, speech recognition, language modelling, acoustic modelling, systematic review

1 Introduction

During the last two years, Covid-19 lock-downs forced universities worldwide to move to the virtual world of online lectures. This situation had many disadvantages that mainly related to the social and psychological aspects of the university experience. However, online learning brought forth some advantages as well. An important one, which we will focus on in this paper, is the availability of recorded lectures online, which in turn leads to the opportunity of developing more accurate algorithms for automatic lecture transcription.

Transcribing lectures into text has a number of advantages, e.g., searchability and accessibility to deaf or hard of hearing students, and students who are non-native in the lecture's language. However, transcribing each lecture manually is usually an unreasonable amount of additional work.

This paper thus concerns the automatic extraction of scripts from lecture recordings in terms of the theoretical composition of a transcription system as well as its practical usability and problems. We will not focus exclusively on English transcription, but rather on techniques applicable regardless of language. We will also not be talking about any practical or ethical concerns that may arise from widespread use of automatic transcription, whatever shape those concerns might take.

After introducing our methodology, we will discuss the most common techniques used in these papers by describing the language and acoustic models used, as well as the training and test data. We will then take a look at the most common evaluation metrics chosen for this task. Finally, we will discuss our overall conclusion and propose future work that should be done in this field.

2 Methodology

The first step we took was assembling 30 papers by searching the keywords ”automatic lecture transcription”, ”automatic video transcription”, ”lecture transcription”, and ”lecture speech transcription”, always using the AND operator. We also alternated the keyword ”automatic” with ”automated”, which yielded different results. We used the Google Scholar, IEEExplore, and Scopus databases to conduct our search.

2.1 Inclusion and exclusion criteria

We decided to include peer-reviewed papers in English, that reported an empirical research of an application used to transcribe lectures into text. We focused only on higher-education lectures, and decided to include only papers that dealt with post-lecture transcription (PLT), as opposed to real-time captioning (RTC) techniques, since the former usually yields better results (Ranchal et al. 2013).

Initially, we decided to include papers that described an application concerning transcribing the English language, our reasoning being that algorithms and language models may differ between languages. However, after reviewing a large number of papers, we noticed that the methods are more or less the same across languages. Therefore, we decided to drop the language criterion.

We searched by most cited papers, keeping in mind the relevance to the specific topic and other inclusion criteria. Table 1 shows a summary of the final inclusion and exclusion criteria.

Table 1: Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
Peer-reviewed journals	Non-peer-reviewed journals
Higher-education lectures	Not higher-education
Empirical primary research	Other types of research (e.g. systematic reviews)
PLT techniques	RTC techniques

2.2 Final included papers

Our final step was to review together the 30 papers that were gathered, and to read the abstracts and overall methods of the papers. After discussing each paper shortly, and according to the aforementioned criteria, we clustered the papers into *relevant*, *irrelevant*, and *somewhat relevant* piles.

The *somewhat relevant* pile included papers that did not meet one or two of our criteria, mainly not being about higher-education or using RTC techniques. We decided to separate those from the *irrelevant* pile because we assumed that they can still be used if needed.

The final *relevant* pile included 11 papers which we decided to read thoroughly and include in our final review. Figure 1 summarizes the overall process described above.

2.3 Limitations

Our methodology does not include papers published in journals that are not written in English, which could result in missing applications made for more languages. Also, it does not consider any papers not published in journals (e.g. book chapters) which could potentially be very relevant to the topic. Therefore, ideally, more research should be done in the future taking these limitations into account.

3 Results

The papers come from various discipline backgrounds of the authors such as computer science, electrical engineering, mechatronics, and communication science. The vast majority of papers (8 of 11) were published at the Institute of Electrical and Electronics Engineering (IEEE) International Conference between 2001 and

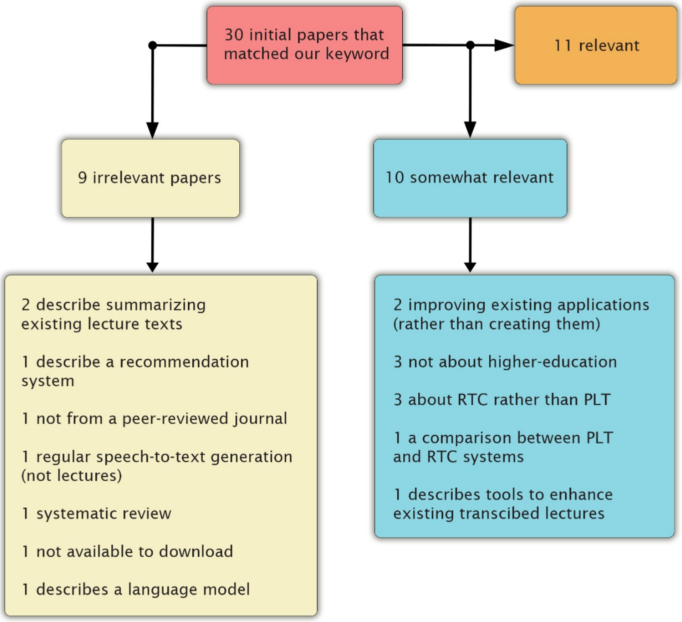


Figure 1: Clustering process of papers

2013, such as the IEEE Workshop on Automatic Speech Recognition and Understanding in 2001, the IEEE International Conference on Acoustics, Speech, and Signal Processing in 2003, and the 12th International Conference on Machine Learning and Applications in 2013. Table 2 shows a summary of languages, countries, and authors' affiliations.

3.1 Models

To create an automatic speech recognition (ASR) system that produces transcripts based on audio recordings of lectures, two models have to work side by side. The first is an acoustic model, which takes as input audio after feature extraction and transforms it to text using the second, a language model. Figure 2 shows a flowchart describing this pipeline.

3.1.1 Acoustic models

All of our papers agree on their choice of Hidden Markov Models (HMMs) as acoustic models. Moreover, those which specified all used triphone states and Gaussian Mixture Models.

Table 2: Languages, countries, and affiliations

Language	Country	Affiliation
Japanese	Japan	Kyoto University
		Tokyo Institute of Technology
		NTT Communication Science Laboratories
	South Africa	University of Stellenbosch
English	South Africa	University of Cape Town
	The Netherlands	University of Twente
	Canada	University of Toronto
	England	University of Sheffield
	Italy	Centro per la Ricerca Scientifica e Tecnologica
German	Germany	University of Potsdam
Spanish	Spain	Universitat Politecnica de Valencia
Czech	Czech Republic	Technical University of Liberec

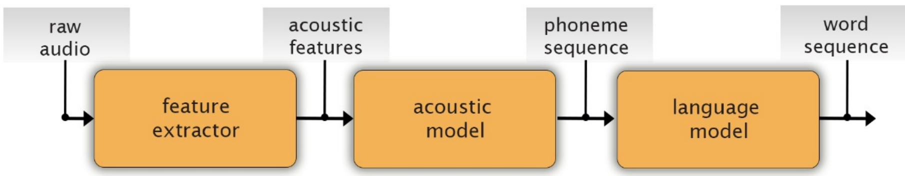


Figure 2: Model pipeline for automatic speech recognition

The acoustic model relates acoustic features extracted from the raw audio data to phonetic sequences.

In a HMM, the acoustic features are seen as the observations caused by hidden states representing phonemes or parts thereof, each state being associated with observation probabilities represented as a mixture of Gaussians, and state transition probabilities represented as a global matrix. Adding the left and right context to each phonetic state (that is, the phonetic units before and after it, since those usually affect pronunciation) produces *triphones*.

This model would be unnecessarily large, since many phonetic sequences produce similar enough acoustic features that they must share many states (e.g. ”back”, ”pack” and ”bag”). This means that model parameters may be re-used between them; this is called *state tying*. Which states are similar enough to be tied together is determined by decision trees using binary linguistic questions (e.g. ”Is the left context a voiced fricative?”).

Decoding for a new utterance is done by finding the most likely path through the combined language and acoustic models, i.e. the string that most likely corresponds to the phonetic state sequence that most likely produced the observed audio features. Exhaustive search through the sequences is rarely possible, as for any real application the state space would be too large. The search space may therefore be reduced by the use of e.g. beam search.

3.1.2 Language models

Again, all our papers agree on using n-gram language models, with a roughly halfway split between trigram and 4-gram models among those which specified. A few papers did not mention what cardinality of n-gram they have used.

Within the transcription system, the language model exists to relate the transcribed string of text to the phonetic output of the acoustic model. To relate pronunciation and spelling, the language model is trained on a phonetic dictionary; these are generally standardized for each language. Many different words and word sequences, however, are pronounced the same way while being spelt very differently (e.g. ”recognize speech” vs. ”wreck a nice beach”). The probability of a certain word occurring depends on the preceding words (e.g. ”speech” being more likely than ”beach” after ”recognize”). In order to not let these dependency chains become unreasonably long, an n-gram model assumes that the likelihood of a word depends only on the n words directly before it.

3.1.3 Model improvements

In the reviewed papers, several improvements to the models were demonstrated.

A common refinement of the transcription process is speaker adaption, done by Constrained Maximum Likelihood Linear Regression (CMLLR), as per Giuliani et al. (2004). As the name implies, a linear transformation learned through regression is applied to the observation probabilities of the HMM. Since these are Gaussians, the transformation affects their mean and variance. This allows the model to filter out the idiosyncrasies of the training speaker or the channel (e.g. a noisy microphone) while preserving the relevant part of the features. One would think that this is detrimental to the ability of the model to generalize to other speakers, but because adaption can be done with only a few seconds of speech, the only downside of applying it for every speaker is increased model complexity.

The language model, too, can be refined: Since it is trained on a closed vocabulary set, it has to deal with words outside of this set (out-of-vocabulary or *OOV words*) by trying to approximate them phonetically, and since lectures are likely to use jargon not included in a general dictionary, it makes sense to enrich the training dictionary with terms related to the domain of the lecture, as suggested in Cerva et al. (2012). These terms may be obtained from pre-existing transcripts or from resources like textbooks.

Another adaption to the language model is not as relevant to English, but more to other languages (e.g. Japanese and Finnish) where there is a significant difference between spoken and written language. Lectures will be given in spoken style, but transcripts should be done in written style rather than verbatim (this applies to English, too, insofar as filler noises and speech errors should be left out). It has been proposed in Akita et al. (2012) to use a Bayesian statistical transformation model to switch between spoken and written styles, so that the language model can be trained on transcriptions approximating spontaneous speech without their peculiarities spreading into the final output.

3.2 Training and testing data

In this section, we will discuss the datasets used to train each model, as well as the test datasets used to evaluate the final transcripts produced by the ASR system.

3.2.1 Training data for acoustic models

In order to train the acoustic models, audio recording of each language should be obtained. Preferably, recordings should be of people with diverse accents, backgrounds, and genders to be able to generalize better across speakers. However, this was not always possible due to the lack of proper resources.

Most papers used available recordings of different audio from various resources, not necessarily of academic lectures. For example, Munteanu et al. (2007) used the Wall Street Journal (WSJ) Dictation Corpus, which holds about 30 hours of speech from 283 speakers. Kawahara et al. (2001) and Kawahara et al. (2008) used the CSJ which has about 700 hours of recordings, some of which from various technical conferences. Papers that used only recordings of academic lectures, e.g. Niesler & Willett (2002), had a notable decrease in the duration of data compared to other papers, using an overall of 38 hours of recorded material.

With the exception of the aforementioned paper, all papers used a training dataset of about 100 hours at least.

3.2.2 Training data for language models

Since our chosen papers deal with a variety of different languages, each one has different approaches in collecting linguistic corpora to train their respective language models.

For low-resource languages like Czech, a corpus had to be collected from different sources that included as much spontaneous speech as possible, as well as specific academic terminology. The corpus thus consisted of, among others, text from bachelor and master theses, and verbatim transcriptions of telephone calls and radio debates (Cerva et al. 2012).

For Japanese systems (Kawahara et al. 2001, Kawahara et al. 2008), the Corpus of Spontaneous Japanese (CSJ) (Furui 2003) was used, which includes 612 presentations and their transcriptions. Another paper that dealt with Japanese used transcriptions of presentations in conferences that were concerned with speech, acoustics, linguistics, and the Japanese language (Niesler & Willett 2002).

Papers concerning Spanish used the poliMedia corpus, which was created by manually transcribing over 700 lectures that correspond to 100 hours of speech in Spanish. The corpus was created as part of the EU-funded project *transLectures*[1], which explores inventive and economical tools for transcribing and translating educational video material (Martínez-Villaronga et al. 2013, Silvestre-Cerdà et al. 2013).

As for German models, Yang et al. (2011) used several general corpora like the Leipzig-Wortschatz, DBPedia, and German daily news, with the addition of transcripts of specific lectures to incorporate the academic terminology aspect into the model.

Finally, with English being the language with the most resources, different corpora were used by different authors. Mbogho & Marquard (2013) used the English Wikipedia because of the broad topics it covers, which include academic terminology. Leeuwis et al. (2003) used the Translanguage English Database (TED), which has 39 transcriptions of lectures. And AlHarbi & Hain (2012) used the Liberated Learning Consortium (LLC) corpus of lectures. The LLC is a network of international researchers, with one of its aims being improving ASR systems of captioning and transcription.

Additionally, some authors decided to add custom-made corpora developed by extracting information from the slides of each lecture and using a web-query of either the bullet points or the keywords from the slides to retrieve relevant documents (Munteanu et al. 2007, Kawahara et al. 2008). These documents were

[1] http://www.translectures.eu

then added to the corpora used to train the model, with the aim of receiving a better accuracy by integrating topic-specific terms.

3.2.3 Testing data

All papers tested their systems on actual recorded lectures and their transcriptions. The number of lectures that were tested on ranged from 4 (Munteanu et al. 2007) to 23 (Martínez-Villaronga et al. 2013), with the average overall duration of lectures being about 3 hours (calculated based on the data available in each paper). When possible, the authors tried to diversify the topics of the lectures being tested on (Cerva et al. 2012, AlHarbi & Hain 2012, Mbogho & Marquard 2013).

3.3 Evaluation metrics

The most commonly used metric to evaluate ASR systems is *Word Error Rate* (WER), which is calculated with the formula:

$$WER = \frac{S + D + I}{N}$$

Where S is the number of word substitutions, D is the number of word deletions, I is the number of word insertions, and N is the number of words in the reference transcript. The lower the WER, the better the model (Yang et al. 2011).

All of the papers we reviewed use WER as a metric of evaluation in some form. 10 papers use it directly, and 1 paper uses it indirectly by calculating *Word Accuracy* which is the complement of WER (i.e. *1 - WER*).

The second most commonly used metric for language models is *perplexity* (PP) utilized by 6 out of 11 papers. This metric measures the probability of the test set, normalized by the number of words. The lower the PP, the better the model (Mbogho & Marquard 2013).

Other metrics used are percentage or number of OOV words, which measures how many words are not in the dictionary of the system. The lower this measure, the better. *Word Correct Rate* (WCR) is another metric used to measure how many words are correctly recognized as a proportion of the total number of words in the reference transcript; this measure is very similar to WER, except it does not take insertions into account (Mbogho & Marquard 2013).

Finally, papers that took keywords into account when building their ASR system measured *precision* (i.e. the sum of true positive keywords, divided by the sum of true positive and false positive keywords), *recall* (i.e. the sum of true positive keywords, divided by the sum of true positive and false negative keywords) (Munteanu et al. 2007), and *F-measure* (i.e. the harmonic mean of precision and recall) (Kawahara et al. 2008).

4 Conclusion

Our main conclusion from these papers is two-fold: For one, even though the results of the papers cannot reasonably be compared to each other due to the use of different corpora for training and testing, they have come to very similar conclusions on which models to use - perhaps this indicates the problem being seen as already solved as best as needed.

More importantly, though, all of the papers are at least eight years old and none of them use deep learning or neural networks, instead relying on Hidden Markov Models and statistical n-gram models. Research on automatic lecture transcription seems to have largely stalled in recent years, despite significant advances in all fields of natural language processing.

We believe more research on applying modern language processing techniques, especially deep learning, to be direly necessary. In particular, this work is necessary to improve the accessibility of recordings to students with hearing impairments.

References

Akita, Y., M. Watanabe & Tatsuya Kawahara. 2012. Automatic transcription of lecture speech using language model based on speaking-style transformation of proceeding texts. *13th Annual Conference of the International Speech Communication Association 2012 (Interspeech 2012)* 3. 2323–2326. DOI: 10.21437/Interspeech.2012-610.

AlHarbi, Ghada & Thomas Hain. 2012. Automatic transcription of academic lectures from diverse disciplines. In *2012 IEEE spoken language technology workshop (SLT)*, 398–403. Miami: IEEE. DOI: 10.1109/SLT.2012.6424257.

Cerva, Petr, Jan Silovsky, Jindrich Zdansky, Ondrej Smola, Karel Blavka, Karel Palecek, Jan Nouza & Jiri Malek. 2012. Browsing, indexing and automatic transcription of lectures for distance learning. In *2012 IEEE 14th international workshop on multimedia signal processing (MMSP)*, 198–202. Banff: IEEE. DOI: 10.1109/MMSP.2012.6343440.

Furui, Sadaoki. 2003. Recent advances in spontaneous speech recognition and understanding. In *ISCA & IEEE workshop on spontaneous speech processing and recognition*, 1–6. Tokyo: ISCA.

Giuliani, Diego, Matteo Gerosa & Fabio Brugnara. 2004. Speaker normalization through constrained MLLR based transforms. In *8th international conference on spoken language processing (Interspeech 2004)*, 2893–2896. Jeju Island: ISCA. DOI: 10.21437/Interspeech.2004-9.

Kawahara, T., H. Nanjo & S. Furui. 2001. Automatic transcription of spontaneous lecture speech. In *IEEE workshop on automatic speech recognition and understanding, 2001. ASRU '01.* 186–189. Madonna di Campiglio: IEEE. DOI: 10.1109/ASRU.2001.1034618.

Kawahara, Tatsuya, Yusuke Nemoto & Yuya Akita. 2008. Automatic lecture transcription by exploiting presentation slide information for language model adaptation. In *2008 IEEE international conference on acoustics, speech and signal processing*, 4929–4932. Las Vegas: IEEE. DOI: 10.1109/ICASSP.2008.4518763.

Leeuwis, E., M. Federico & M. Cettolo. 2003. Language modeling and transcription of the TED corpus lectures. In *2003 IEEE international conference on acoustics, speech, and signal processing, 2003. proceedings. (ICASSP '03).* Vol. 1, 228–232. Hong Kong: IEEE. DOI: 10.1109/ICASSP.2003.1198760.

Martínez-Villaronga, Adrià, Miguel A. del Agua, Jesús Andrés-Ferrer & Alfons Juan. 2013. Language model adaptation for video lectures transcription. In *2013 IEEE international conference on acoustics, speech and signal processing*, 8450–8454. Vancouver: IEEE. DOI: 10.1109/ICASSP.2013.6639314.

Mbogho, Audrey & Stephen Marquard. 2013. Improving the transcription of academic lectures for information retrieval. In *2013 12th international conference on machine learning and applications*, vol. 2, 560–567. Miami: IEEE. DOI: 10.1109/ICMLA.2013.177.

Munteanu, Cosmin, Gerald Penn & Ronald Baecker. 2007. Web-based language modeling for automatic lecture transcription. In *8th annual conference of the international speech communication association*, 2353–2356. Antwerp: ISCA. DOI: 10.21437/Interspeech.2007-266.

Niesler, Thomas & Daniel Willett. 2002. Unsupervised language model adaptation for lecture speech transcription. In John H.L. Hansen & Bryan L. Pellom (eds.), vol. 2, 1413–1416. Denver: ISCA.

Ranchal, Rohit, Teresa Taber-Doughty, Yiren Guo, Keith Bain, Heather Martin, J. Paul Robinson & Bradley S. Duerstock. 2013. Using speech recognition for real-time captioning and lecture transcription in the classroom. *IEEE Transactions on Learning Technologies* 6(4). 299–311. DOI: 10.1109/TLT.2013.21.

Silvestre-Cerdà, Joan Albert, Alejandro Pérez, Manuel Jiménez, Carlos Turró, Alfons Juan & Jorge Civera. 2013. A system architecture to support cost-effective transcription and translation of large video lecture repositories. In *2013 IEEE international conference on systems, man, and cybernetics*, 3994–3999. Manchester: IEEE. DOI: 10.1109/SMC.2013.682.

Yang, Haojin, Christoph Oehlke & Christoph Meinel. 2011. German speech recognition: a solution for the analysis and processing of lecture recordings. In *2011 10th IEEE/ACIS international conference on computer and information science*, 201–206. Sanya: IEEE. DOI: 10.1109/ICIS.2011.38.

Chapter 11

Knowledge representation approaches in adaptive educational AI systems

Kristina Sigetova, Eliasz Ganning & Tobi Obeck

The vast amount of available data in today's world makes it hard to find the right information. Adaptive systems approach this problem by recommending content based on the users' preferences. In a similar manner, in the field of artificial intelligence in education, adaptive systems also play an important role. It is common for such systems to build a model not only based on the learner but also on what shall be learned (domain model). We decided to give an overview of specifically domain knowledge representation approaches. Therefore, we conducted a systematic literature search in SCOPUS that aimed to find out the current state of the research field. A well-defined search query with applied filter criteria yielded 21 articles. These were analyzed and synthesized on common topics, showing a dominating interest in recommender systems and ontologies. Some of the most common types of recommender systems are presented and discussed regarding knowledge representation. Our results indicate that the research field is moving towards solutions for the cold start problem and two-part frameworks that combine AI-methods with knowledge representation.

Keywords: knowledge representation, ontology, adaptive learning, recommender systems, knowledge-based filtering

1 Introduction

Nowadays, there is an abundance of media available for informing oneself about a specific topic or domain of interest. This media can range from comprehensive and in-depth books or video courses to short-form content like disconnected blog articles or tweets. For a newcomer to a domain, it is not obvious what information is the most relevant to know at the beginning, nor how different concepts relate to each other.

The authors of this paper have experienced this overwhelming feeling themselves while familiarizing themselves with the study of Cognitive Science - a broad domain encompassing a number of areas that can also stand alone, such as Artificial Intelligence, Psychology, Neuroscience, Linguistics and Philosophy.

Miller (2018) describes knowledge maps that make entering a new field more accessible and less cognitively demanding. Knowledge maps communicate the structure of knowledge and the relationship of concepts in a visual way. They are most effective when accompanied by multiple summaries written in different levels of detail. His statements indicate that a visual presentation of knowledge can be beneficial for learners.

This visual approach for learning made us curious how knowledge is presented, recommended and represented in adaptive computer systems used in education. Especially, since Ivanova (2021) states that modern educational systems increasingly make use of artificial intelligence methods, e.g. for automating the admissions or assessment process, but also the more apparent ones, such as intelligent tutoring systems (ITS) or adaptive learning systems.

The focus of this paper is to find out which approaches have been implemented within educational systems that address the problem of complex knowledge representation. Specifically, we wanted to see how content is represented in a way that is accessible to a computer system in order to improve the experience of learners. Gutierrez & Sequeda (2021) explained, that the inspiration comes from ancient visual representations of knowledge, such as knowledge graphs and taxonomies. The combination of knowledge and data has been an increasingly exciting development in computer science in recent decades.

It is also worth noting that not only domain knowledge, but also other components are often part of a knowledge representation model. These include the learner model and the pedagogical model, where the latter encompasses the decisions or rules that guide the teaching path. We decided to put our main focus on educational systems that utilize artificial intelligence to adapt towards the learners' needs and interests and thereby create a more individual and effective learning experience.

Before conducting the actual systematic literature review, previous similar work has been regarded to get an initial overview of the field. One particularly salient study by Hatzilygeroudis & Prentzas (2006) carried out a comprehensive review. They provided this list of requirements for a satisfactory knowledge representation:

- Domain expert: domain knowledge is supplied by an expert, who delivers not only content, but also relationships between the different concepts; likewise, they assist with validation checks of the resulting model.

- Naturalness of representation: the less abstract and technical the model is, the easier it is for knowledge engineers to implement and also update it in the future.

- Knowledge characteristics: the main distinctions are structural (hierarchical) and relational knowledge. The authors also mention heuristic representation, however, only the first two are relevant for domain knowledge. A knowledge model represents the main concepts within the subject, how they relate to one another, i.e. relational, as well as the corresponding prerequisites or specializations, i.e. hierarchical (such as course modules).

According to the review by Hatzilygeroudis & Prentzas (2006), the main knowledge representation types comprise structured representations, such as semantic nets, symbolic rules (if-then), case-based representations, neural networks, probabilistic nets and hybrid systems, such as neurofuzzy representations or descriptive logics. Since the review is almost two decades old, we wanted to carry out a new review, focusing on more recent papers. Our aim is that this paper serves as a good overview on the most current approaches for knowledge representation and can highlight trends when seen in comparison to Hatzilygeroudis & Prentzas (2006).

The remaining part of our review is organized as follows: First, we provide a thorough description of our literature search methodology. Then, we will introduce our main findings, starting with a quantitative summary and next move through the most common themes that we have identified. Finally, we will critically look at what we found as well as our approach to the searching process. We end by providing some concluding remarks.

2 Methodology

To get an understanding of the field and to prepare a search string, some initial research was carried out. By looking into articles and literature reviews, we gained insights into possible search terms. This initial research also laid the ground for the inclusion and exclusion criteria.

A first draft of the search string was tested to find out how promising the results were. After looking into the top results of the first draft search hits, it

Table 1: The final search string

”knowledge representation” OR ”Ontology” OR ”ontologies”	W/10
”educational” OR ”e-learning” OR ”adaptive learning”	W/10
”AI” OR ”artificial intelligence” OR ”fuzzy logic” OR ”neural networks” OR ”Bayesian network” OR ”Genetic algorithms”	

was concluded that the majority of the articles treated the subject of knowledge representation as a side topic. Some search words were added and the search string was also modified to include the W/10 operator (two strings must be within a ten words distance from each other) as seen in Table 1. This did not only provide more accurate search hits but also reduced the amount of search hits down to a manageable size of 155 articles in total. The first draft search string also helped with developing the inclusion and exclusion criteria which can be seen in Table 2. Both the first draft search string and the final search string were passed into the advanced search of SCOPUS[1].

Table 2: Inclusion and exclusion criteria

Inclusion criteria:	Exclusion criteria:
• Document type: journal or conference proceedings • Final publication • Publication year: 2016-2021 • Written in English	• Not using AI technology • Outside of the learner/education domain

This method further consisted of a two step filtering process as seen in Figure 1. The first filtering was carried out by looking at the article title, keywords and abstract as well as evaluating relevance in relation to the exclusion criteria. This resulted in including 47 of the 155 articles found on SCOPUS. Before continuing the filtering, some more inclusion and exclusion criteria were specified such as: Excluding articles without domain knowledge representation and excluding articles theorizing about knowledge representation. The second filtering consisted of looking into the whole article and finding parts that fit the updated inclusion and exclusion criteria. After deleting duplicates or inaccessible articles the final

[1]See more at: https://www.scopus.com/search/form.uri?display=advanced (institutional access required)

filtering yielded 21 articles. All articles were randomized in their order between filtering phases and randomly divided between team members to minimize possible review bias.

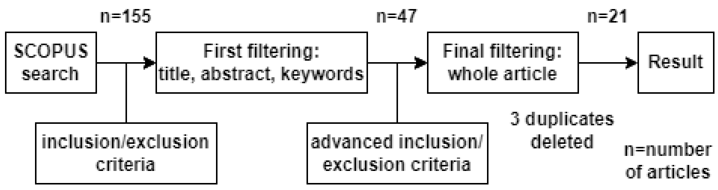


Figure 1: Filtering process

The resulting 21 papers were subsequently analysed in order to identify the main themes that addressed knowledge representation approaches.

3 Results

In the following section, we are going to present a summary of the common themes, interesting remarks and general notions of our literature sample.

3.1 Introduction of the Results

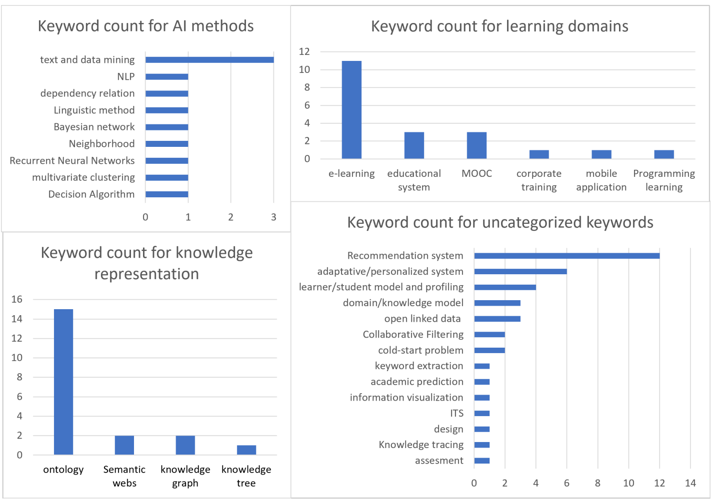


Figure 2: Bar charts representing the keyword count for the four different categories

To gain quantitative insight into the topics mentioned in the final 21 articles, keywords from the articles' metadata were extracted and counted. Keywords addressing the same concepts were regarded as synonyms and thus counted as the same; keywords assessed as redundant or not descriptive enough were removed, such as *competencies* or *ACM/IEEE computing curriculum*. To get a better overview of the results, a categorization of the keywords was carried out. Four distinct categories were identified: AI methods, learning domains, knowledge representation and uncategorized keywords as seen in Figure 2. Across all categories, the five most frequent keywords were: *ontology* (15), *recommender system* (12), *e-learning* (11), *adaptive/personalized system* (6), and *learner/student model and profiling* (4).

3.2 Recommender Systems in Education

One approach that helps users find the information most relevant to them is the use of recommender systems (RS), which also proved to be the most common adaptive system in the literature that we examined. Apart from education, they are commonly used in e-commerce to recommend products or on social media platforms to suggest content.

In educational systems, such as platforms for e-learning or massive open online courses (MOOCs), RSs are used to provide learners with recommendations for content from the web (e.g. Aeiad & Meziane 2019, Gulzar et al. 2019, Ibrahim & Yang 2019), courses or modules to take (Bakanova et al. 2019, Gulzar et al. 2018, Harrathi et al. 2018), or exercises to test their acquired knowledge (Diao et al. 2018, Wu et al. 2020). Likewise they also assist teachers with developing their curriculum based on the students' needs (Sebbaq et al. 2020). However, special care must be taken with RSs in an educational context (Wu et al. 2020). It is necessary to store the prior knowledge of the learner and the knowledge of concepts the learner is supposed to learn. This storage consists of a student model and a domain model.

3.3 Types of Recommender Systems and the Cold Start Problem

There are several approaches to implement a RS Bhareti et al. (2020). We briefly present a selection of these approaches that we found most frequently in our results. This overview provides helpful background information for less experienced readers. We cover collaborative filtering (CF), content-based filtering (CBF), hybrid approaches, and knowledge-based filtering (KBF). The following explanation is mostly based on Rocca (2019) and Bhareti et al. (2020).

CF works by counting the interaction of users with items; for example users rating movies or whether products were bought. These interactions are stored in an interaction matrix. There are two approaches to determine recommendations. Firstly, the user-user approach makes item recommendations by finding users with similar preferences and suggesting new items that these similar users liked. Secondly, the item-item approach takes items that the user already likes and then searches for similar items by referring to the interaction history of other users with an overlap of their liked items. This implementation of a CF is called memory-based, but another popular approach is the model-based one. Here machine learning techniques are applied to represent the user-item interaction with a model, e.g. by using matrix factorization to create a dense matrix representation. Such a model can then be used to give new suggestions. However, a downside in the model-based approach is the resulting harder interpretation of the matrix compared to the memory-based approach.

A common problem that arises when using CF is the cold start problem which occurs when only few interactions are captured. This is common when new items or users enter the system. In this case, a CF system does not have enough information to make good recommendations. CBF systems suffer far less from this problem. They take characteristics about the user and the item into account; for example the age, gender, location and education of a user and the category, actors, and duration of a movie. Since this kind of data should usually be available even for new users and items, a proper recommendation can be determined. To accomplish this, CBFs use classification or regression algorithms to create a model, which then predicts whether items will be liked.

There are two other popular ways to mitigate the cold start problem. The first one is hybrid systems which combine multiple approaches in one. According to Bhareti et al. (2020),"[...] a hybrid RS always gives superior results over the conventional algorithms [...]". Another approach is KBF which is especially useful when the user behavior could not have been observed yet and only few ratings are present. Instead user preferences and interests are taken into account which usually need to be queried first. The cold start problem seems to be a common problem for adaptive systems in education because we observed that a number of papers put emphasis on it (Jeevamol & Renumol 2021, Joy et al. 2021, Mbaye 2018, Tarus et al. 2017).

3.4 Knowledge-Based Filtering and Ontologies

Our results included all of the approaches outlined above, often combining them. However, the most relevant type of a RS for domain knowledge representation is KBF. KBF systems are oftentimes used in conjunction with ontologies – a database-like structure that defines relations between different concepts and entities – to model the domain knowledge.

Several studies have addressed the construction of ontologies. Chimalakonda & Nori (2020) list three main ways of developing ontologies: manually by domain experts, by means of a semi-automated process that makes use of techniques such as text mining or natural language processing, or fully automated methods. Moreover, Bakanova et al. (2019) pointed out three prerequisites for implementing an ontology: setting the conceptual boundaries of the system to be developed, a list of the necessary content items to be covered by the representation, and making sure that the chosen RS is able to work with it. As a result, the developed ontology serves as a framework of the domain. A common methodology for the ontology-based systems was to first search for existing repositories and then adjusting them to one's needs. Examples of such approaches are Chimalakonda & Nori (2020) and Andaloussi et al. (2017). A popular open source tool to edit ontologies is *Protégé* which is explicitly named within their methods by e.g. Shishehchi et al. (2021) and Ibrahim & Yang (2019).

3.5 Further Results

As mentioned, it is not uncommon to combine several approaches in the development of a knowledge model. In our sample, these included a semantic web by Carbonaro (2021), knowledge graphs by Carbonaro (2021), integrating CF and a decision algorithm with ontologies by Mbaye (2018) and sequential pattern mining by Tarus et al. (2017). A hybrid system consisting of ontology, knowledge and rules has been implemented by Harrathi et al. (2018).

Among the systems which aim to provide recommendations from web-based learning resources, NLP techniques using text-mining (Aeiad & Meziane 2019), n-grams (Gulzar et al. 2019) or WordNet (Gulzar et al. 2018) have been found useful to identify topics or suitable keywords for resource filtering. Another commonality that we found was that domain knowledge representation was not implemented as a separate module within a system, but rather interconnected with learner's knowledge representation (Jeevamol & Renumol 2021, Sebbaq et al. 2020, Shishehchi et al. 2021, Tarus et al. 2017) or pedagogical methods (Carbonaro 2021, Harrathi & Braham 2021). Chimalakonda & Nori (2020) go as far as to include instructional knowledge, such as learning context, goals, process, evaluation and environment alongside the content within their knowledge domain.

Some advantages and disadvantages of ontologies for knowledge representation have been a point for discussion. Numerous authors agree that the flexibility as well as ability to represent complex domains of knowledge make them a powerful technique for domain modelling (Andaloussi et al. 2017, Harrathi & Braham 2021), and that their generalizability makes them shareable and re-usable (Shishehchi et al. 2021, Chimalakonda & Nori 2020). On the other hand, Andaloussi et al. (2017) criticize that they are an expensive and complex method for obtaining a knowledge representation. Due to the high abstraction characteristic of ontologies, field experts are a necessary part of the creation process in order to validate the resulting framework.

Apart from ontologies, we have also identified a number of other methods of knowledge representation within our sample. Diao et al. (2018) used a handcrafted course knowledge tree to represent the learning points of different exercises and relationships between them. In contrast, Wu et al. (2020) decided to mathematically represent the individual exercises used in the learning course by the distribution of different knowledge concepts within them, using a long short-term memory neural network to evaluate the knowledge concept coverage of an exercise. Joy et al. (2021) mentioned that they made use of a combination of semantic processes, however, they do not provide further information that would have helped us to better understand their approach.

Two studies also provided a short overview of other methods they had identified. First, Andaloussi et al. (2017) reported that semantic web was the most common one, along with ontologies, metadata and machine learning techniques, also confirming our observation that the methods are indeed often combined. Second, Harrathi & Braham (2021) identified fuzzy cognitive maps and vector-based modeling which show an overlap with the earlier review of Hatzilygeroudis & Prentzas (2006), but also remarking that the large majority of literature does not specify their domain modeling approach in sufficient detail.

Finally, where provided, we noted the discipline, for which the adaptive learning system had been designed and were initially surprised to find that the field of choice tended to be computer science or programming curriculum (e.g. Tarus et al. 2017, Diao et al. 2018, Gulzar et al. 2019), with two notable exceptions: the mobile app of Bakanova et al. (2019) which focused on training and educational tutorials for employees within information and communication technologies; and an e-learning system designed for adult literacy teaching across different languages (Chimalakonda & Nori 2020). It is likely that the subject of knowledge representation was chosen by the authors due to convenience, since they were likely to be already familiar with its structure due to their own educational background.

4 Conclusion

Although the results provided some interesting insight to the field of knowledge representation in education, there seems to be a difference in this review's findings and similar reviews of the field. For instance, there are a lot more different types of AI methods used with knowledge representation, such as those previously mentioned in Hatzilygeroudis & Prentzas (2006).

It is also interesting that topics on ontologies and RSs were much more common than other topics. That the results are skewed towards these topics could be explained by the following reasons: these topics are recently more popular and research for the last few years has been dominated by them, or that concepts such as RSs are strongly connected to knowledge representation in general.

Another more critical reason could be that the review's search string presented an isolated part of the field. For instance, one could argue that the exclusion of the search terms ”ontology” and ”ontologies”, or the inclusion of other knowledge representation related words such as ”semantic webs”, ”learning objects”, ”data mining”, or ”domain model” could have resulted in a more balanced representation of the field. However, even though previous iterations of the search string did present other methods of knowledge representation, they also presented both an overwhelming amount of search hits and a lot of completely unrelated articles.

In conclusion, our findings show that the field of AI-driven knowledge representation in education is strongly connected to RSs that present, through ontology based knowledge representation, adaptable domain models. AI methods such as CF, CBF, hybrid approaches, and KBF are commonly used and sometimes combined together with knowledge representations within a joint framework. Most prominent interests in the current research address the cold start problem by using hybrid approaches in order to optimize RSs.

References

Aeiad, Eiman & Farid Meziane. 2019. An adaptable and personalised E-learning system applied to computer science Programmes design. en. *Education and Information Technologies* 24(2). 1485–1509. DOI: 10.1007/s10639-018-9836-x. https://doi.org/10.1007/s10639-018-9836-x (20 December, 2021).

Andaloussi, Kenza Sakout, Laurence Capus & Ismail Berrada. 2017. Adaptive Educational Hypermedia Systems: Current Developments and Challenges. In *Proceedings of the 2nd international Conference on Big Data, Cloud and Applications* (BDCA'17), 1–8. New York, NY, USA: Association for Computing Machinery. DOI: 10.1145/3090354.3090448. https://doi.org/10.1145/3090354.3090448 (20 December, 2021).

Bakanova, A., N.E. Letov, D. Kaibassova, K.S. Kuzmin, K.V. Loginov & A.N. Shikov. 2019. The use of ontologies in the development of a mobile e-learning application in the process of staff adaptation. English. *International Journal of Recent Technology and Engineering* 8(2 Special Issue 10). 780–789. DOI: 10.35940/ijrte.B1144.0982S1019.

Bhareti, K., S. Perera, S. Jamal, M.H. Pallege, V. Akash & S. Wiieweera. 2020. A Literature Review of Recommendation Systems. English. In DOI: 10.1109/INOCON50539.2020.9298450.

Carbonaro, Antonella. 2021. Modelling educational resources to support accessibility requirements. English. In DOI: 10.1109/CCNC49032.2021.9369554.

Chimalakonda, Sridhar & Kesav Nori. 2020. An ontology based modeling framework for design of educational technologies. English. *Smart Learning Environments* 7(1). DOI: 10.1186/s40561-020-00135-6.

Diao, Xiuli, Qingtian Zeng, Hua Duan, Faming Lu & Changhong Zhou. 2018. Personalized exercise recommendation driven by learning objective within e-learning systems. English. *International Journal of Performability Engineering* 14(10). 2532–2544. DOI: 10.23940/ijpe.18.10.p29.25322544.

Gulzar, Zameer, Anny Leema & Gerard Deepak. 2018. PCRS: Personalized Course Recommender System Based on Hybrid Approach. English. In vol. 125, 518–524. ISSN: 1877-0509. DOI: 10.1016/j.procs.2017.12.067.

Gulzar, Zameer, Arun Raj & Anny Leema. 2019. Ontology supported hybrid recommender system with threshold based nearest neighbourhood approach. English. *International Journal of Information and Communication Technology Education* 15(2). 85–107. DOI: 10.4018/IJICTE.2019040106.

Gutierrez, Claudio & Juan F. Sequeda. 2021. Knowledge graphs. *Communications of the ACM* 64(3). 96–104. DOI: 10.1145/3418294. https://doi.org/10.1145/3418294 (14 December, 2021).

Harrathi, M. & R. Braham. 2021. Recommenders in improving students' engagement in large scale open learning. English. In vol. 192, 1121–1131. ISSN: 1877-0509. DOI: 10.1016/j.procs.2021.08.115.

Harrathi, M., N. Touzani & R. Braham. 2018. A hybrid knowlegde-based approach for recommending massive learning activities. English. In vol. 2017-October, 49–54. ISSN: 2161-5322. DOI: 10.1109/AICCSA.2017.150.

Hatzilygeroudis, Ioannis & Jim Prentzas. 2006. Knowledge Representation in Intelligent Educational Systems. In *Web-based Intelligent E-learning Systems: Technologies and Applications*, 175–192. Journal Abbreviation: Web-based Intelligent E-learning Systems: Technologies and Applications. DOI: 10.4018/978-1-59140-729-4.

Ibrahim, M. & Y. Yang. 2019. An ontology-based web crawling approach for the retrieval of materials in the educational domain. English. In vol. 2, 900–906. DOI: 10.5220/0007692009000906.

Ivanova, T. 2021. Modern technologies and e-learning - Present and future. English. In 165–170. DOI: 10.1145/3472410.3472423.

Jeevamol, J. & V.G. Renumol. 2021. An ontology-based hybrid e-learning content recommender system for alleviating the cold-start problem. English. *Education and Information Technologies* 26(4). 4993–5022. DOI: 10.1007/s10639-021-10508-0.

Joy, J., N.S. Raj & V.G. Renumol. 2021. Ontology-based E-learning Content Recommender System for Addressing the Pure Cold-start Problem. English. *Journal of Data and Information Quality* 13(3). DOI: 10.1145/3429251.

Mbaye, B. 2018. Recommender system: Collaborative filtering of e-learning resources. English. In vol. 2018-July, 213–217.

Miller, Francis. 2018. Organising knowledge with multi-level content: Making knowledge easier to understand, remember and communicate. en. *Organising Knowledge*. 91.

Rocca, Baptiste. 2019. *Introduction to recommender systems*. en. https://towardsdatascience.com/introduction-to-recommender-systems-6c66cf15ada (6 January, 2022).

Sebbaq, Hanane, Nour Eddine El Faddouli & Samir Bennani. 2020. Recommender system to support moocs teachers: Framework based on ontology and linked data. English. In 103–108. ISSN: 2153-1633. DOI: 10.1145/3419604.3419619.

Shishehchi, S., N.A.M. Zin & E.A.A. Seman. 2021. Ontology-Based Recommender System for a Learning Sequence in Programming Languages. English. *International Journal of Emerging Technologies in Learning* 16(12). 123–141. DOI: 10.3991/ijet.v16i12.21451.

Tarus, J.K., Z. Niu & A. Yousif. 2017. A hybrid knowledge-based recommender system for e-learning based on ontology and sequential pattern mining. English. *Future Generation Computer Systems* 72. 37–48. DOI: 10.1016/j.future.2017.02.049.

Wu, Z., M. Li, Y. Tang & Q. Liang. 2020. Exercise recommendation based on knowledge concept prediction. English. *Knowledge-Based Systems* 210. DOI: 10.1016/j.knosys.2020.106481.

Part III

Extended Reality and Robots

Chapter 12

Combining AI and virtual reality in medical education

Jakob Lohkamp, Felix Naujoks & Juri Moriße

In this chapter, a literature review with focus on the integration of novel medical training paradigms was conducted. It includes several papers which discuss the implications of a virtually simulated surgical training, communication training, but also assessment and evaluation of the aforementioned. Results of individual papers suggest that deploying artificial intelligence (AI) applications in combination with virtual reality (VR) or augmented reality (AR) can advance medical training not only for students but also for educators. It is also deemed beneficial for the development of medical applications as it allows to gather new data that provides the opportunity to gain new insights. However, we see a few shortcomings of the studies at hand that are due to small and biased sample sizes, ill-conducted implementation and in-accountability of user experience. While these concerns should be addressed in future research, the mixture of AI and VR/AR seems immensely promising in the field of medical education. It not only complements existing teaching formats, but enriches them in a practical and efficient manner.

Keywords: Artificial Intelligence, Virtual Reality, Education, Simulation, Medicine

1 Introduction

Rapid advancements in AI give rise to new possibilities in virtually all scientific fields and enable the development, as well as enhancement of an array of technologies. The effects increasingly extend to the field of education. The subfield of medical education has a long history of educational practices that are relying on personal interaction between students, educators, and patients. There is continuous adaption to both medical and technological developments. Incorporating the ever-changing advancements of technology into medical practice requires not only state-of-the-art training for students but also for practicing physicians. Case-based or problem-based learning is an emerging concept in teaching where traditional lectures are reduced to focus on student contact and self-directed learning (Bligh 1995). We see high potential in combined approaches to medical education where AI-applications are embedded in a simulation case-based learning paradigm. This has several reasons. Firstly, medical education has long foreseen the benefits of such case-based learning, and it is widely integrated into the standard curriculum. Secondly, cost of resources in a high-quality educational program such as medicine is high as students are relying on close personal mentoring and guidance by experts. Additionally, lots of information has to be conveyed in limited time, furthermore increasing the need for efficiency. For a more in-depth analysis see Bligh (1995) for an overview on medical education paradigms and challenges. We think that AI in combination with VR/AR has the potential to fill these needs, since it can be widely deployed, is potentially cost and time efficient once incorporated, and aids existing teaching structures by individualization and evaluation of students progress. For this review, we propose that educational methods that combine VR/AR applications with AI provide an opportunity to manifest and support new forms of learning. In the following we will investigate this hypothesis by answering a number of questions. We will begin by illustrating the methodology according to which we searched and selected papers for this review and will list the questions according to which we analysed the papers. Furthermore, we try to answer the proposed questions, and finally, we come back to discuss our initial hypothesis about the improvements that intelligent VR/AR applications can yield in medical education.

2 Methodology

Our review is based upon the pubmed database, the U.S. national library of medicine by the National Centre of Biotechnology Information, in which we used the following search string: “artificial intelligence virtual reality education simulation”. As we want to examine the current state of development we included articles only published 2019 or later. Our initial search yielded 63 papers. We minimized our search space even further and extended our search string by ”NOT review” to exclude reviews and meta-analyses, which led to 44 articles in total. We then narrowed our search to studies conducted with AI and some form of simulation (VR/AR). Of the 44 remaining articles 33 had to be excluded and 3 were inaccessible (14 were not concerned with education, 11 did not include any form of AI, 3 were review papers, 3 were out of context, 2 did not include VR). With the remaining corpus of 8 articles we set to answer the following five questions:

1. What are the author's motivations for investigating the use of intelligent VR/AR simulations in medical education?
2. For what medical domains were intelligent VR/AR or simulation techniques presented?
3. What methods of AI were deployed?
4. What are the papers' results?
5. Do the papers discuss possible ethical problems?

3 Results

3.1 What are the author's motivations for investigating the use of intelligent VR/AR simulations in medical education?

The papers at hand differ widely in their motivation. The overlapping categories can be described as (A) expertise assessment and metrics identification (B) introducing new technologies to aid medical training, and lastly (C) facilitation of transparency in AI.

Mirchi et al. (2019) use artifical neural networks to investigate expertise in a discectomy task. Performance is also assessed by Bissonnette et al. (2019) and Siyar et al. (2020). Being able to identify expertise performance has great benefits for teachers as well as students. As implicit knowledge can be captured in multidimensional feature spaces, spaces that are humanely in-conceivable, but an AI can excel in outlining very specific factors (Mirchi et al. 2019, Siyar et al. 2020).

In order to achieve the aforementioned some authors propose the development of new, more objective metrics that make expertise measurable (Bissonnette et al. 2019) or are specifically tailored for VR-Simulation tasks (Mirchi et al. 2019). Paysan et al. (2021) identify a need for surgical activities to be labelled automatically to further enhance AI-applications in gathering more sensor- and motion based video data for training purposes. By doing so, AI-Applications can reduce the need for tedious labeling process researchers are facing right now.

Melnyk et al. (2021) provide a new method for improving the teaching in a robotic suturing task. They investigate the advantage to novice students obtained when augmenting such VR surgery task with information about visual gaze patterns expert surgeons exhibit during such a task.

Additionally, linking expertise classification, objective feedback based on metrics, and instructor input creates a novel educational tool which in return allows for automation of traditional forms of teaching while re-defining educational goals (Mirchi et al. 2020, Maicher et al. 2019).

Shorey et al. (2020) specifically looked into communication training for nursery students by the application of virtual patients. They aimed to investigate not only the students' but also the educators' attitudes towards virtual patients. Mirchi et al. (2020) aimed to increase transparency of AI in a medical context, simultaneously proposing and validating a new framework for deploying explainable AI in a medical simulation training context.

3.2 For what medical domains were intelligent VR/AR or simulation techniques presented?

Surgery is the main field presented in our literature (6/8). Inside surgery the application domains vary from brain surgery (Siyar et al. 2020) and neck surgery (Bissonnette et al. 2019) to robotic surgical training (Melnyk et al. 2021), see Table 1 for a summary. Another area is the usage of virtual patients in the subdomains of nursing or information gathering, i.e. taking a patients medical history (Shorey et al. 2020, Maicher et al. 2019).

One of the aims of simulated surgical tasks is to define objective metrics for the assessment of medical students (Mirchi et al. 2019, 2020). Therefore, most studies have a goal of developing methods that can be generalized to many different areas of surgery, i.e. that the current possibilities of surgical simulation will include many specific medical domains which are not mentioned in Table 1.

Table 1: Medical domains the papers deal with.

surgery	virtual patients
subpial brain tumor resection task	nursing
anterior cervical discectomy	information gathering
hemilaminectomy	
cerebral tumor procedures	
robotic surgical training	
surgical activity recognition	

3.3 What methods of AI were deployed?

In this section we will only include a brief overview of the methods used in our corpus, for a more detailed explanation of AI applications and procedures we advise the reader to look up Russell & Norvig (2010).

Although, we have a rather homogeneous distribution of medical domains, very different choices were made when it comes to the applied methods (see Table 2). Several authors make use of Artificial Neuronal Networks (ANN) and Convolutional Neuronal Networks (CNN), but also other supervised machine learning methods in the form of Support Vector Machines (SVM) are used. A couple of studies mix different methods to compare their performance but also combine the best performing applications. Unfortunately, two papers, Melnyk et al. (2021) and Shorey et al. (2020), did not include detailed information on what methods were deployed.

Table 2: AI methods used in the papers.

study	methods
Mirchi et al. (2020)	supervised ML, ANNs, SVM
Mirchi et al. (2019)	supervised ML, ANNs
Bissonnette et al. (2019)	SVM, LDA, KNN, Naive Bayes, Decision Tree
Shorey et al. (2020)	n.a. speech recognition
Siyar et al. (2020)	KNN, KDE, SVM, FKNN
Melnyk et al. (2021)	Eye tracker by pupil lab
Paysan et al. (2021)	ANNs, CNN, MLP, semi-Markov model
Maicher et al. (2019)	NLP (ChatScript), logistic regression, CNN

Bissonnette et al. (2019) identified SVMs as the best performing application, with an accuracy as high as 97.6% of correct classifications of experts. Mirchi et al. (2020) developed a SVM that classifies participants on 4 metrics with an accuracy as high as 92% of correct classifications of skilled and novice participants.

Siyar et al. (2020) applied a Fuzzy K-Nearest Neighbor algorithm with overall error rates as low as 8.3%, as their best performing classifier.

Paysan et al. (2021) and Mirchi et al. (2019) used Artificial Neural Networks for classification, where Paysan et al. (2021) focused on Convolutional Neural Networks and a Multi Layer Perceptron. Mirchi et al. (2019) implemented an ANN for classifying 3 different levels of surgical expertise with a training accuracy of 100% and a testing accuracy of 83.3%.

In contrast, Maicher et al. (2019) made use of Natural Language Processing (NLP) software ChatScript. ChatScript is primarily a pattern matching system that is well suited to doctor–patient interactions.

3.4 What are the papers' results?

We identified three distinct categories for dividing our papers and will present the results separately for each category. We will start with the group of papers that investigate possible further improvements for teaching methods that already make use of VR applications. This is also the category to which most papers in this review belong. We then go on to present the results of the second category in which alternatives to traditional teaching methods are introduced. This category is comprised of only one paper. Lastly, we will present the results of the third category of papers in which the relationship between applications and users is investigated. For this last category our review also includes only one paper.

Regarding the improvement of already existing VR teaching methods, Melnyk et al. (2021) found that augmenting a VR robotic suturing task with additional information of expert gaze behaviour is beneficial for learning. Compared to augmenting the training process only with information on the movement of experts, the 'expert' information helped students to complete the task more efficiently. The authors note that the advantage of gaze-related augmented training is strongest in the early stages of learning.

Bissonnette et al. (2019), Siyar et al. (2020), Mirchi et al. (2020), Mirchi et al. (2019) and Paysan et al. (2021) are all, in part, addressing the same problem of evaluating surgical performance by identifying metrics and features that indicate a trainee's expertise. The papers differ in regard to the surgical task they introduce an evaluation system for (see Table 1). Bissonnette et al. (2019) found twelve metrics that relate to the trainee's performance in a VR-based hemilaminectomy (a spinal surgery task). Using these metrics, the authors built a Support Vector Machine that was able to achieve a 97.6% accuracy in differentiating between senior and junior surgeons. Mirchi et al. (2020) also designed a system intended to identify features that can be used to classify students as skilled or novice. The surgical task was a simulated VR brain tumor resection. Using only 4 metrics, the authors systems achieved a 92% accuracy in the classification task. Siyar et al. (2020) introduced a system that also is concerned with a brain tumor resection task. Here, the authors identified 15 features that allow a support vector machine to achieve a 90% accuracy in the classification task. Mirchi et al. (2019) had intentions as the previous papers but shifted their focus on a VR-based anterior cervical discectomy. Several performance metrics were identified and additionally weighed with regards to their importance for assessing a surgeons performance in this task. While Bissonnette et al. (2019), Mirchi et al. (2020) and Mirchi et al. (2019) are all investigating ways to improve the evaluation of VR-based surgery performance via identifying performance metrics and relations among those, they also argue these metrics could be used in a traditional education context. For example, Mirchi et al. (2019) propose that identifying the weighting of performance metrics that optimizes the classification performance of a Support Vector Machine bears important information also for traditional teaching. The reasoning behind this is that assigning different importance to different aspects of performance in a surgical task allows for a more holistic evaluation. While the paper by Paysan et al. (2021) is also intending to improve already existing applications, it has a different aim. The focus here lies on developing a system for surgical activity recognition that relies less on expert annotations than previous systems. The advantage of such an approach is a lesser need for labeling of surgical activity which is time and cost intensive process. While the authors argue that their presented system could improve surgical activity recognition with less expert annotation, they also state that "[...] high intersample variance, the small sample size and the lack of sufficient activity annotations of the used data set do not allow for more general statements"(Paysan et al. 2021: p. 6).

For the second category, Maicher et al. (2019) is introducing a system that can be seen as a direct alternative to traditional education methods, advocating the use of virtual patients to teach students about the acquirement of medical histories. For evaluation, the authors compare the systems' evaluation of student performance to several human raters. They find that the system's evaluation performance is comparable to that of human raters and therefore conclude that the system is a viable alternative to traditional methods for teaching medical history acquisition.

Lastly, the paper of Shorey et al. (2020) neither tries to improve traditional educational methods nor AI related methods but instead investigates students' and educators' attitudes towards virtual patients like the ones introduced in Maicher et al. (2019). According to the authors, the use of virtual patients is potentially an effective tool for teaching nursing communication skills. For the authors virtual patients seem to be a good additional tool that is time and cost effective and also provides a good benchmark for students' performance. However, they also note that using virtual patients lacks authenticity and is drastically hindered by other limitations such as dysfunctional speech recognition which can lead to frustration and poor conversation flow. They conclude that the limitations should be solved before an official implementation.

Overall, it can be said that all papers produced promising results for advocating the use of intelligent VR applications in medical education.

3.5 Do the papers discuss possible ethical problems?

Our review showed that most publications are not discussing any ethical problems of their research (6/8), or include demographic information of their participants (5/8). This section focuses on two publications by Mirchi et al. (2020, 2019), where ethical problems are discussed and demographic information is provided. The following ethical concerns apply to all surgical simulation tasks, if they assess and evaluate student performance on objective/standardised metrics.

Biased data is a typical ethical problem in AI research (Ntoutsi et al. 2020), as the applied AI methods in our literature are heavily data dependent they are prone to be biased. The two publications by Mirchi et al. (2020) indicate a possible gender bias. In Mirchi et al. (2020) only 18% of the participants were female and in Mirchi et al. (2019) only 14,29% were female see Table 3.

It is questionable if gender differences determine different approaches to surgical tasks, this should be investigated. But, at the current state of research, gender biased data should not be generalised without clearly indicating the limits of sample size and demographic backgrounds, as it does not suffice for an objective assessment nor an objective evaluation of student performance. Additionally, as sample sizes are generally low in our literature and the VR/AR simulation platforms, such as the Sim-Ortho platform, are exclusively developed for right-handed participants (Mirchi et al. 2019) the generalisation of findings might only apply to a certain subset of participants and further studies with a broader demographic background and more participants should be conducted.

Table 3: Participants split by gender

study	participants	male	female
Mirchi et al. (2020)	50	41	9
Mirchi et al. (2019)	21	18	3
Shorey et al. (2020)	24	6	18
Bissonnette et al. (2019)	41	n.a.	n.a.
Siyar et al. (2020)	115	n.a.	n.a.
Melnyk et al. (2021)	17	n.a.	n.a.
Paysan et al. (2021)	10	n.a.	n.a.
Maicher et al. (2019)	102	n.a.	n.a.

Another typical point of discussion is the transparency of artificial neural networks (Mirchi et al. 2020). If a classification of a student is done by an artificial neural network, but the underlying reasons for this classification are not explainable, students and teachers loose trust in the classification (Mirchi et al. 2020). If students and teacher loose trust, AI is not able to assert itself and would fail to enhance current educational methods.

Another discussed ethical problem is the illusion of being skilled (Mirchi et al. 2020), if learning is standardised students may learn how to ”cheat” the algorithm, meaning that they find a way to change their classification without learning the actual underlying (psychomotor) skill they should learn. In the context of medicine this could have the ethical implication that the falsely classified student is a risk to the patient as the necessary skills are not acquired.

An often ignored ethical concern of standardised student assessment and evaluation is the inability to replicate the affective component of feedback (Mirchi et al. 2020). Neither current cognitive state nor the emotions of the learner can be assessed by the shown methods. The learner can therefore feel disconnected and the received feedback might not be accepted, which prohibits the learning experience and declines skill acquirement (Mirchi et al. 2020).

4 Conclusion

In conclusion, the overall opportunities of AI-VR/AR systems for medical education seem to be very promising. As we have seen, the different papers achieve their best results with a variety of methods which indicates that the appropriate selection of a method for a specific task is of great importance. In addition to that, many of the papers make use of supervised machine learning methods which rely on data that bears valuable information for training. Given the design of a system and the usability of a data set usually are specific to only one or a small number of applications, we see a broad use of AI methods in general medical education reliant on further research that focuses on making systems more general and less reliant on labelled data. From the papers discussed in this review, only Paysan et al. (2021) tackle this problem.

The biggest potential we see in the great efficiency that well-designed AI/VR methods can bring. In light of medical school’s notoriously packed schedules, parallel theoretical and practical training, and the general long time of education, AI/VR systems can in our opinion contribute to a better and more individualized learning experience for students as well as a reduced workload for teaching staff.

However, we also see some limitations that still need to be addressed before intelligent VR/AR systems can be commonly deployed in medical education. This is also supported by some of the papers that indicate that an optimal implementation is crucial for success. If the process is rushed or ill-prepared it has the effect of worsening the learner’s experience, as it was the case in Shorey et al.

(2020) where the speech recognition was poorly designed and led to user frustration. We want to highlight that only very few studies completely specified demographics and sample size of student participants and no paper discussed the diversity of used simulations. This may be the result of a lack of awareness for ethical concerns such as gender, race and other biases in general or the authors may simply argue that their systems are not prone to any bias or that resources have been a limiting factor regarding sample size and diversity. Nevertheless, we think ethical concerns should at least have been discussed and their implications for medical education should be further investigated to prevent the potential development of biases that may bring a disadvantage to either certain students or patients.

Still, we believe that VR in combination with AI has great potential for educators as well as students to further improve medical training. Through individualization, augmentation of VR with helpful information, instant feedback and constant supply of generated training exercises students' learning experience can be enhanced. On the other hand, better identification and understanding of performance metrics as well as a reduced workload due to automated exercise generation and evaluation of training exercises are in the interest of educators to further improve the quality of their teaching. We argue that this is generally supported by the scientific field as the heavy focus on improving already existing methods for VR-based education methods may indicate that research is already past the process of establishing that VR methods are a valuable addition to medical education and proceeded to further improve these approaches with the help of AI tools.

Taking these points into consideration, we conclude that most intelligent VR/AR applications are not suited to fully replace some traditional teaching methods yet, but rather provide a tool to facilitate understanding and evaluation of performance in simulated robotic surgery tasks. However, for specific tasks, e.g. taking patients medical history, we have seen that systems can already provide ways of learning that bring advantages to both students and educators while maintaining the quality medical education requires. In the long term, we believe that intelligent VR/AR systems will be able to provide such support to a wide variety of fields in medical education and will be a valuable addition and sometimes even substitution of traditional teaching methods.

References

Bissonnette, V, N Mirchi, N Ledwos, G Alsidieri & RF Winkler-Schwartz A Del Maestro. 2019. Artificial intelligence distinguishes surgical training levels in a virtual reality spinal task. *Journal of Bone and Joint Surgery, American Volume* 101(23). e127 1–8. DOI: 10.2106/JBJS.18.01197.

Bligh, John. 1995. Problem-based learning in medicine: an introduction. *Postgraduate Medical Journal* 71(836). 323–326. DOI: 10.1136/pgmj.71.836.323. https://pmj.bmj.com/content/71/836/323.

Maicher, Kellen R., Laura Zimmerman, Bruce Wilcox, Beth Liston, Holly Cronau, Allison Macerollo, Lifeng Jin, Evan Jaffe, Michael White, Eric Fosler-Lussier, William Schuler, David P. Way & Douglas R. Danforth. 2019. Using virtual standardized patients to accurately assess information gathering skills in medical students. *Medical Teacher* 41(9). 1053–1059. DOI: 10.1080/0142159X.2019.1616683.

Melnyk, Rachel, Timothy Campbell, Tyler Holler, Katherine Cameron, Patrick Saba, Michael William Witthaus, Jean Joseph & Ahmed Ghazi. 2021. See like an expert: gaze-augmented training enhances skill acquisition in a virtual reality robotic suturing task. *Journal of Endourology* 35(3). 376–382. DOI: 10.1089/end.2020.0445.

Mirchi, Nykan, Vincent Bissonnette, Nicole Ledwos, Alexander Winkler-Schwartz, Recai Yilmaz, Bekir Karlik & Rolando F Del Maestro. 2019. Artificial neural networks to assess virtual reality anterior cervical discectomy performance. *Operative Neurosurgery* 19(1). 65–75. DOI: 10.1093/ons/opz359.

Mirchi, Nykan, Vincent Bissonnette, R Yilmaz, N Ledwos, Alexander Winkler-Schwartz & Rolando F Del Maestro. 2020. The virtual operative assistant: an explainable artificial intelligence tool for simulation-based training in surgery and medicine. *PLOS*. 15.

Ntoutsi, E., P. Fafalios, U. Gadiraju, V. Iosifidis, W. Nejdl, M. Vidal, S. Ruggieri, F. Turini, S. Papadopoulos, E. Krasanakis, I. Kompatsiaris & K. Kinder-Kurlanda. 2020. Bias in data-driven artificial intelligence systems - an introductory survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 10(3). 3.

Paysan, Daniel, Luis Haug, Michael Bajka, Markus Oelhafen & Joachim M. Buhmann. 2021. Self-supervised representation learning for surgical activity recognition. *International Journal of Computer Assisted Radiology and Surgery* 16(11). 2037–2044. DOI: 10.1007/s11548-021-02493-z.

Russell, S. & Peter Norvig. 2010. *Artificial intelligence, a modern approach. fourth edition.*

Shorey, Shefaly, Emily Ang, Esperanza Debby NG, John Yap, Lydia Siew Tiang Lau & Chee Kong Chui. 2020. Communication skills training using virtual reality: a descriptive qualitative study. *Nurse Education Today* 94. DOI: 10.1016/j.nedt.2020.104592.

Siyar, Samaneh, Hamed Azarnoush, Saeid Rashidi, Alexander Winkler-Schwartz, Vincent Bisonnette, Nirros Ponnudurai & Rolando F Del Maestro. 2020. Machine learning distinguishes neurosurgical skill levels in a virtual reality tumor resection task. *Medical and Biological Engineering and Computing* 58. 1357–1367. DOI: 10.1007/s11517-020-02155-3.

Chapter 13

Making science fiction reality - applications, benefits and drawbacks of 3D holograms in education

Henriette Kohnen, Alina Ohnesorge, Henriette Uhlenbrock & Dennis Witowski

Abstract: The use of interactive 3D holograms is one of many promising novel applications of Artificial Intelligence in education. While the technology is not yet widely spread in educational institutions, chances are that this will change in the near future. This paper gives an overview on ways to use interactive holograms in education and learning based on a literature review and presents the potential advantages and disadvantages of their application. It becomes clear that interactive 3D holograms can be used at many different levels and in various fields of the educational sector, reaching from primary schools to higher education. Holograms cannot only be used as an aid to explain specific items to a group of students, but they can also function as teaching assistants to help students with impairments. The research proved that many of those applications are highly useful in education and learning in more than one way, while there are also some downsides. Those mainly concern the missing digital infrastructure and the costs of well-functioning and useful hologram applications, which in most cases are still too high.

Keywords: hologram, 3D hologram, digital hologram, education, learning, interactive learning

1 Introduction

The future of education is inevitably linked with the development of new technologies. With recent advances in the field of Artificial Intelligence (AI), new opportunities for education arise. The interest in AI has been growing in the recent years, and AI in education has been researched for the past 30 years. Even though digitalization and the application of new technologies are progressing slowly, particularly in education, experimental use of new technology can already be observed in many places. See for example Huang et al. (2016), who explored the usage of virtual reality in tourism. Educators are just starting to discover the wide range of potential opportunities that AI applications can provide for learning and education (Zawacki-Richter et al. 2019), one of them being (interactive) holograms.

The term hologram originates in the Greek words *holo*, which means 'full vision', and *gram*, meaning 'written'. Holograms emerge from the positive interference of light waves that create the illusion of three-dimensional images when recorded. The term holography then refers to the whole record of all the information about the form and the audiovisual resources of the projected object (Sancho et al. 2019).

Holograms are not a new technology. Their roots date back to the mid-19th century when English engineer Henry Dircks pioneered the technology he called Dircksian Phantasmogaria. A few years later, John Henry Pepper adapted the technology for his famous Pepper's Ghost illusion. To achieve the desired effect of a ghostly image on stage, a clear sheet of glass is installed at a certain angle in relation to an actor located on a lower level underneath the stage, as shown in Figure 1. Modern holography was invented in the 1940s by Dennis Gabor, a Hungarian physicist who was awarded the Nobel Prize in 1971. The technology was implemented in the 1960s when laser technology had advanced far enough. Today, holograms are famously used to bring deceased singers or celebrities that have not been real in the first place on stage: Tupac performed at Coachella 2012, Michael Jackson at the 2014 Billboard Music Awards, and Hatsune Miku is a virtual K-pop star (Kaufman 2017).

The practical realizations of holograms vary widely. While Hatsune Miku, the above-mentioned singer is realized as an advanced version of Pepper's Ghost, holograms as an optical reality in contrast to an illusion are much less popular. Interactive holograms are created where the optical illusion of the interfering light waves is combined with modern AI and machine learning technologies. Patel & Bhalodiya (2019) suggest to add natural language processing units as well as a camera to enable interaction with the 3D object. Israeli start-up RealView is among the first companies to have developed an interactive live hologram which has successfully been used in clinical medical imaging (Bruckheimer et al. 2016).

To us, interactive holograms seem to be a particularly promising application

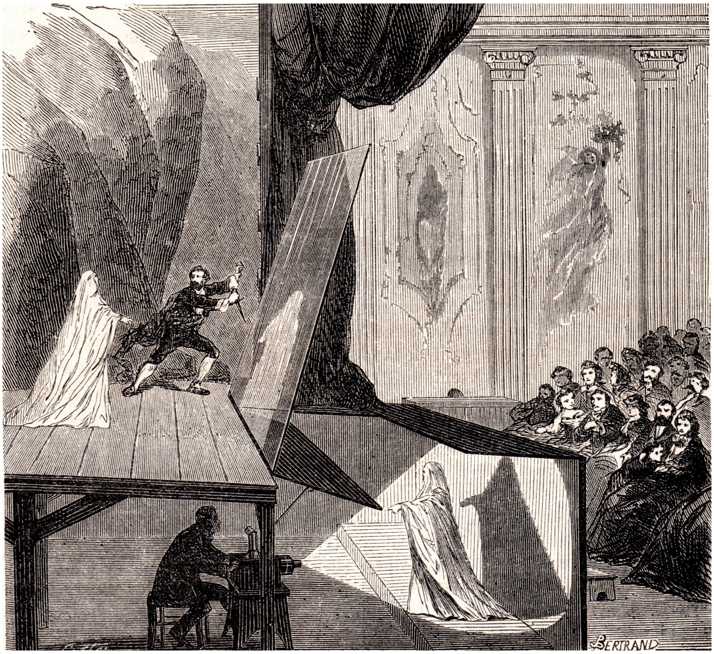


Figure 1: Pepper's Ghost stage set up, ©Le Monde Illustré (1862), Public Domain

of technology in education as they provide a wide range of very diverse opportunities and have proven to be helpful for a variety of tasks.

Similar to technologies like virtual reality (VR) and augmented reality (AR), holograms have only recently found their way into education. As the technology is still in its infancy, we decided to not limit our research solely to interactive 3D holograms similar to those developed by RealView. Instead, we also included non-interactive technology for our second research question, while focusing on interactive holograms for the first research question. As a result, holograms overlap with the AR and VR technologies. In many cases, additional tools like glasses or reflective surfaces are necessary to see the hologram (see for example Adamo-Villani & Anasingaraju (2016), Mavrikios et al. (2019)).

Consequently, our goal is to examine the connection between the field of education and the technology of holograms. To that end, we conducted a literature review and examined two research questions:

- How are interactive holograms used in education and learning?

- What are the advantages and disadvantages of the application of holograms in the educational field?

2 Methodology

In order to investigate our research questions presented above, we included 21 papers in our literature review and focused on 15 of these papers.

2.1 Data acquisition

Data was collected between November 2021 and January 2022. In order to find suitable papers, we used the search engines Google Scholar (https://scholar.google.com/) and Semantic Scholar (https://www.semanticscholar.org/). In our search, we primarily combined the keywords *hologram* and *education*. This approach initially proved very fruitful but yielded too broad a spectrum of papers. Therefore, we decided to additionally use the keywords *interactive hologram*, combined *hologram* and *interaction*, replaced the keyword *education* with *learning*, and added *AI* to each of these search strings.

2.2 Data corpus

Our initial data corpus contained 28 papers. In order to assess if all of these were suitable, we read the papers' abstracts and skimmed through the main parts. For a paper to make it into our final data corpus, it had to a) connect the topics of education and holograms and b) be relevant to our research questions and hence include application fields and/or advantages and disadvantages of the use of such holograms. In addition, several papers not meeting these criteria were used for the introductory chapter of this paper. A mandatory criterion for all papers was that they had to stem from reputable journals or conferences and were peer-reviewed. In the end, we excluded seven papers completely, while six were found not to be relevant in all aspects, but still useful for individual references. The main focus of our analysis is on the remaining 15 papers. Out of the 21 included papers, a total of seven papers were not cited directly as information was often similar across several papers.

2.3 Analysis

All papers we decided to use were thoroughly read by at least one of us. Papers were read by additional members of our research group when further clarification was needed. While reading, important parts were marked and collected in another document. This collection of notes forms the basis for this paper.

3 Results

The analysis of our 21 papers yielded numerous applications, advantages and disadvantages.

3.1 Application of interactive holograms in education and learning

Applications of (interactive) holograms are already found in several fields and at different levels of education, reaching from primary schools to higher education institutions.

Sancho et al. (2019), for example, provide an idea for a low-budget realization of holograms that function as an extension rather than a temporal replacement of the teacher present in the classroom: When a teacher is reading to the class, holograms depict in real-time what they are saying, which helps the development of primary school children's reading competence. In that specific case, the holograms are created by a combination of the open-source software Blender, a programme called Ni-Mate, which is already sufficient in its free version, and a sensor called Kinect which can among other things detect the teacher's movements.

However, this is only a very small niche in the wide field of applications: In general, holograms could provide a more interactive and interesting learning experience by using a variety of 3D models (Patel & Bhalodiya 2019) and integrating holograms into already more frequently used technologies like video conferences (Luévano et al. 2015). In contrast to the common form of long-distance communication via video calls where the conversation partners see each other on screens, holograms can be used as a more convincing and realistic mode of distance communication between students and teachers (Kalansooriya et al. 2015). They can also bring famous characters back to life (Kalansooriya et al. 2015), which might be of special interest for history lessons, allowing witnesses of highly important historic events, for example holocaust survivors, to speak to the students.

A larger number of applications feature in the field of medical education. Here, 3D hologram technology is, for instance, used to replace traditional textbooks (Hackett & Proctor 2018) or to simulate a surgery procedure with the help of e.g. the EchoPixel True 3D or the Holo Patient technology (Ramachandiran et al. 2019). Barkhaya & Abd Halim (2016) mention this specifically for the observation and practice of cardio-thoracic surgery as well as neurosurgery where holograms can additionally be used to help educate patients about upcoming procedures. In the field of pharmacology, holograms can display the effect of drugs on (human) internal organs. Further utilization of AI and VR in a medical education context is discussed in Chapter III.

Additional applications in higher education are, for example, found in the fields of archaeology, engineering, and architecture. Here, artifacts or structures that would be far too big, expensive, or fragile to be taken to a traditional classroom or examined in real-life can be studied in great detail and from all possible angles (Barkhaya & Abd Halim 2016, Ramachandiran et al. 2019).

Another application of holograms with a somewhat different link to education is the holographic (AR) signing avatar for deaf students. These avatars translate what the teacher is saying in real-time, and the deaf students wearing special glasses can follow the lecture by watching the avatar located next to the teacher in the classroom (Adamo-Villani & Anasingaraju 2016).

The practical realization of all those different types of holograms can be effected via many different applications that are complicated and expensive in varying degrees: one modality invented by Patel & Bhalodiya (2019) is based on the old technology of the Pepper's Ghost illusion for the generation of 3D holographic objects. They facilitate interaction with the 3D object by adding natural language processing units and a camera, so users can use their voice and gestures.

Another way to visualize holograms is by using passive 3D glasses enabling learners to see virtual objects which they can rotate, grab, or zoom in on and out off, as shown in Figure 2 (Mavrikios et al. 2019).

An even more advanced approach is the so called e-REAL lab, where learners are surrounded by interactive edugraphics as 2D and 3D images, holograms and movies, where no 3D goggles or other tools are necessary as the projections are realized on big screens, walls, and ceilings and as floating 3D holograms (Salvetti & Bertagni 2016).

3.2 Advantages and disadvantages of the application of holograms in the field of education

As shown, holograms have a high potential for use in the field of education. Compared with traditional media, they allow for a more immersive experience since they appear more real than their 2D counterparts. Holograms make communication over long distances possible in a novel way and create new opportunities for the integration of non-living characters into the real world. They specifically

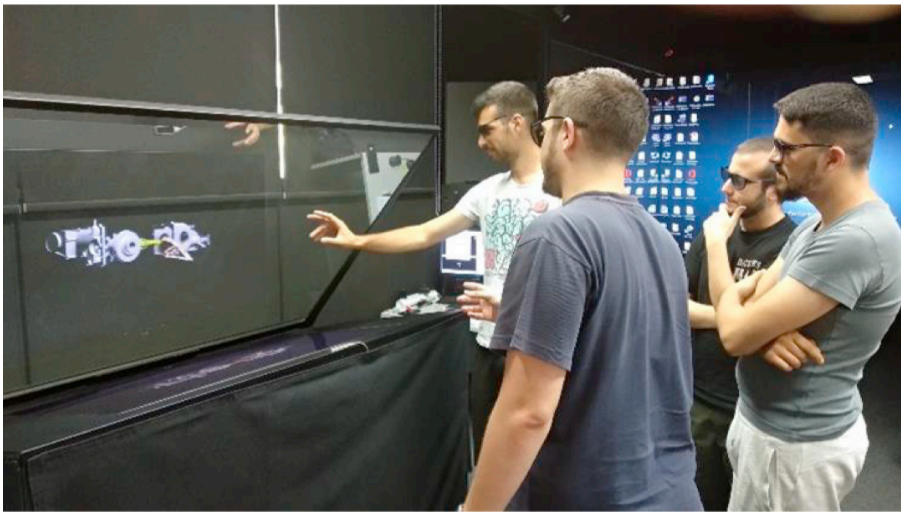


Figure 2: Interacting with a hologram via hand gestures wearing passive 3D glasses (Mavrikios et al. 2019)

create possibilities for structuring knowledge and a better and simplified representation of situations, facts and time-dependent processes, e.g. in the human body (Turk & Seckin Kapucu 2021). Holograms allow for objects to be taken to the classroom that would otherwise would be inaccessible to many students due to their size, fragility or locality (Barkhaya & Abd Halim 2016, Ramachandiran et al. 2019).

Holograms have generally been shown to be a positive experience for students. It has been demonstrated that students deliver better results, absorb a greater amount of knowledge and can access content better if it is taught by using holographic techniques. Students sensed the presence of the holographic teacher similar to them being there in person (Luévano et al. 2015) and the real-time experience had a positive impact on theoretical and practical education (Kalansooriya et al. 2015). In addition, just like virtual environments and the use of virtual agents, holograms may be a future solution to provide personalized learning experiences to students that usually would be unable to afford a personal (human) teacher (Kalansooriya et al. 2015).

While distance learning may provide great opportunities during exigent circumstances, it nevertheless still poses some difficulties to overcome. Compared to traditional ways of communication, holograms have already become a great tool to undo many of the disadvantages of communication over distances (Nadila et al. 2021). However, advanced hologram technology is extremely costly and even more expensive where the necessary infrastructure is not yet available. In particular, a very fast and stable internet connection is required. Usually, holograms can only be viewed from certain angles, which might limit their use in front of big audiences. Often, the individual representations are programmed beforehand as this is an easier technical realization, rendering the resulting hologram not truly interactive (Ramachandiran et al. 2019).

When used, holograms or other visualization techniques can lead to attention problems. The subjects might concentrate on things that are not necessary, thus not accomplishing the learning process to the desired extent (Holland 2019). In order to make hologram technology accessible to the general public, many aspects need to be simplified to make it easier and more intuitive to use (Holland 2019). By way of conclusion, high costs and complex technology can be considered as the biggest challenges for the widespread implementation of holograms.

3.3 Further results

It needs to be stressed how useful holograms can be across the field of education. Turk & Seckin Kapucu (2021) found that the use of holograms is mostly strongly wanted among students, in particular. It is noteworthy that students who perform better academically also have a more positive attitude towards the use of holograms in their education compared with lower-performing students. Nevertheless, the students in this study saw many advantages in using holograms for their learning. They noted, for example, that holograms make the learning more interesting and fun because of their realistic appearance. Unfortunately, holograms are not yet used in most typical classrooms, even though the students in the above study wish for them to be used in all their classes, especially in science classes. While these findings are certainly interesting, it needs to be emphasized that the type of hologram used in this particular study was not interactive and just a simple pyramid reflection.

From a teacher's point of view, too, using holograms in a classroom seems to be beneficial for multiple reasons. Kalansooriya et al. (2015) found that teachers consider holograms a useful tool for distance education because they seem to attract the students' attention and are also expected to help students overcome visualization problems: As the hologram is already in 3D, this might result in an improvement of students' visualization skills. This could enhance their interest in the STEM (science, technology, engineering, mathematics) subjects (Sancho et al. 2019).

4 Discussion

As presented in the sections above, holograms are already being used, have been developed up to a quite considerable degree and seem to be embraced by most students and teachers. However, this far, hardly any student has been able to count holograms as part of their everyday educational life. While holograms are definitely more than just a nice add-on or toy to play around with during class, especially for students suffering from impairments, like the deaf students in Adamo-Villani & Anasingaraju (2016), the disadvantages concerning financial resource problems described as a key factor above remain crucial.

As Turk & Seckin Kapucu (2021) have shown, students were generally very enthusiastic about the use of holograms. Interestingly, the type of hologram used in their study was very small and non-interactive, but a 3D projection on a pyramid made from clear plastic. Arguably, if students are already responding in such a positive manner to small and non-interactive holograms, their enthusiasm for interactive holograms may be even bigger. Additionally, such pyramid holograms are very cheap and can actually be made in class as they only require clear plastic sheets and a phone. Instructions and projection videos are freely available on the internet. This may therefore be a great option until interactive holograms become more affordable.

As presented above, the studies have found that students and teachers show a very positive attitude regarding the use of holograms in class. It may be speculated that these results are due to holograms being a new and exciting technology and the enthusiasm might decrease over time, especially among the students. No long-term studies have yet been conducted, so that this question is open for discussion and further research.

Furthermore, it is also very likely that holographic technology will become cheaper, as it becomes a more reliable and widely used application. This has evidently been the case with many other technologies. Commonly known examples include the development of computers and mobile phones or even novel technologies like VR headsets. a While the idea of VR goes back to the mid-20th century (Flynt 2019), more than 30 years passed between the foundation of VPL Research, Inc. (Burkeman 2001), the first company to sell VR goggles, and the now famous Kickstarter campaign of the Oculus Rift Goggles in 2012 (Oculus 2012).

An increasing amount of research is being carried out in academic institutions, technology companies and the military. Here, holograms are, for example, used for route planning, with additional uses being investigated (Hackett 2013).

5 Conclusion

Researchers have developed numerous promising approaches to bring holography into the classroom and support students' learning processes at different levels and in various fields of their education. Nevertheless, significant challenges remain to be overcome before this technology is ready, with the costs being one of the key problems as they are far too high to use this technology in a typical classroom. One may conclude that there are many educational advantages, but also numerous, mainly economical, disadvantages.

Although most of the approaches are far from ready, the current state of development is very promising and further research needs to be pursued so that one day all students will be able to benefit from this technological advancement in education.

References

Adamo-Villani, Nicoletta & Saikiran Anasingaraju. 2016. Holographic signing avatars for deaf education. *E-Learning, E-Education, and Online-Training*. 54–61. DOI: 10.1007/978-3-319-49625-2_7.

Barkhaya, Nurul Maziah Mohd & Noor Dayana Abd Halim. 2016. A review of application of 3d hologram in education: a meta-analysis. *IEEE 8th International Conference on Engineering Education (ICEED)*. 257–260. DOI: 10.1109/ICEED.2016.7856083.

Bruckheimer, Elchanan, Carmel Rotschild, Tamir Dagan, Gabriel Amir, Aviad Kaufman, Shaul Gelman & Einat Birk. 2016. Computer-generated real-time digital holography: first time use in clinical medical imaging. *European Heart Journal - Cardiovascular Imaging* 17(8). 845–849. DOI: 10.1093/ehjci/jew087.

Burkeman, Oliver. 2001. *The virtual visionary*. https://www.theguardian.com/technology/2001/dec/29/games.academicexperts (13 January, 2022).

Flynt, Joseph. 2019. *The history of vr: when was it created and who invented it?* https://3dinsider.com/vr-history/ (13 January, 2022).

Hackett, Matthew. 2013. *Medical holography for basic anatomy training*. Tech. rep. ARMY RESEARCH LAB ORLANDO FL.

Hackett, Matthew & Michael Proctor. 2018. The effect of autostereoscopic holograms on anatomical knowledge: a randomised trial. *Medical Education* 52(11). 1147–1155. DOI: 10.1111/medu.13729. https://onlinelibrary.wiley.com/doi/abs/10.1111/medu.13729.

Holland, Janet. 2019. Augmented reality: 3d holograms for engaged learning. In *X International Conference The Future of Education: Conference Proceedings. Florença.*

Huang, Yu Chih, Kenneth Frank Backman, Sheila J Backman & Lan Lan Chang. 2016. Exploring the implications of virtual reality technology in tourism marketing: an integrated research framework. *International Journal of Tourism Research* 18(2). 116–128. DOI: https://doi.org/10.1002/jtr.2038.

Kalansooriya, Pradeep, Ashu Marasinghe & K. M. D. N. Bandara. 2015. Assessing the applicability of 3d holographic technology as an enhanced technology for distance learning. *IAFOR Journal of Education* spec ed. 43–57.

Kaufman, Gil. 2017. *Tupac, Michael Jackson, Gorillaz & More: A History of the Musical Hologram.* https://www.billboard.com/music/pop/musical-holograms-history-dead-7717042/ (4 January, 2022).

Le Monde Illustré. 1862. *Pepper's ghost.* https://commons.wikimedia.org/wiki/File:Peppers_Ghost.jpg (13 January, 2022).

Luévano, Eduardo, Eduardo López de Lara & Juan Edward Castro. 2015. Use of telepresence and holographic projection mobile device for college degree level. *Procedia Computer Science* 75. 2015 International Conference Virtual and Augmented Reality in Education, 339–347. DOI: 10.1016/j.procs.2015.12.256.

Mavrikios, Dimitris, Kosmas Alexopoulos, Konstantinos Georgoulias, Sotiris Makris, George Michalos & George Chryssolouris. 2019. Using holograms for visualizing and interacting with educational content in a teaching factory. *Procedia Manufacturing* 31. 404–410. DOI: 10.1016/j.promfg.2019.03.063.

Nadila, Nadila, Mega Wirda Angriani, Diah Novrianty Utami & Dewi Nopita. 2021. Smart three-dimensional (3d) hologram as an innovative teaching tool in virtual learning environment during exigent circumstances. *Pancaran Pendidikan FKIP* 10. 1–14. DOI: 10.25037/pancaran.v10i4.362.

Oculus. 2012. *Oculus Rift: Step Into the Game.* https://www.kickstarter.com/projects/1523379957/oculus-rift-step-into-the-game (13 January, 2022).

Patel, Dhyey & Parth Bhalodiya. 2019. 3d holographic and interactive artificial intelligence system. *2019 International Conference on Smart Systems and Inventive Technology (ICSSIT).* 657–662. DOI: 10.1109/ICSSIT46314.2019.8987926.

Ramachandiran, Chandra Reka, Mien May Chong & Preethi Subramanian. 2019. 3d hologram in futuristic classroom: a review. *Periodicals of Engineering and Natural Sciences (PEN)* 7(2). 580–586. DOI: 10.21533/pen.v7i2.441.

Salvetti, Fernando & Barbara Bertagni. 2016. Interactive holograms and tutorials in healthcare education: case studies from the e-real® experience. *International Journal of Advanced Corporate Learning* 9(2). 13–19. DOI: 10.3991/ijac.v9i2.5988.

Sancho, Julia Vera, Keyda De la Gala Quispe & Klinge Villalba Condori. 2019. Storytelling with holograms for the development of reading competence. In Pablo H. Ruiz & Vanessa Agredo-Delgado (eds.), *Human-Computer Interaction*, 345–355. Corporación Universitaria Comfacauca: Springer. DOI: 10.1007/978-3-030-37386-3_25.

Turk, Hanne & Munise Seckin Kapucu. 2021. Innovative technology applications in science education: digital holography. *Journal of Education in Science, Environment and Health* 7(2). 156–170. DOI: 10.21891/jeseh.748662.

Zawacki-Richter, Olaf, Victoria I. Marín, Melissa Bond & Franziska Gouverneur. 2019. Systematic review of research on artificial intelligence applications in higher education–where are the educators? *International Journal of Educational Technology in Higher Education* 16. 39. DOI: 10.1186/s41239-019-0171-0.

Chapter 14

Humanoid robots in education

Robin Gratz, Lina Lazik, Sönke Lülf & Elisa Palme

With rapid development in the fields of artificial intelligence, including natural language processing, and robotics over the last years, the possibilities for combining the technologies and applying the result in the field of education have increased. To give an overview, we collected examples of prototypes that are already in use. Results show that the experiments conducted with them span across all subjects and the entire age-range associated with institutional learning, from children up to adults. Where the robots mostly act in the form of teachers, tutors or peers, though the boundaries are often fluid. In this the positive effect of anthropomorphism on the learning experience can be observed. Though this has to be handled with care, as upcoming concerns about unhealthy relationships with robots show that there is still further research to be done before a broad application is possible.

Keywords: Humanoid robots, Humanization, Robot tutoring

1 Introduction

"Artificial intelligence (AI) is understood as a type of algorithms or computerized systems that try to imitate to some degree a person's mental processes with their decision-making, making it inherent to study the promising applications of AI in education" (Cóndor-Herrera et al. 2021: p. 184). However, many people working in education are not aware of the advantages AI can already offer and put into prospect through further research.

In general AI in the educational sector can be used in order to support educators, for instance by automatizing teachers' duties, such as correcting tests, or in the optimization of administrative tasks, such as scheduling timetables of both teachers and students. AI further offers students a more suited study experience with regard to strengths and weaknesses in learning. This could, for instance, happen by collecting students' data and analyzing them in order to create smart learning content suited for the specific needs of each individual.

To take the possibilities offered by AI in the field of education even further one can combine smart technologies with devices incorporating technology such as Virtual Assistants (VA). They can be defined as "intelligent agents capable of perceiving their environment, processing perceptions, and responding or acting rationally accordingly" (Cóndor-Herrera et al. 2021: p. 186). Examples for VA are smart speakers such as Siri or Echo. Being based on AI they are able to perform different activities and can imitate the human way of communication by answering in speech rather than simple text output, which may create a sense of talking to a being rather than a machine. However, they are still lacking the humanoid figure or appearance but are rather a shapeless voice in space.

Already incorporating VA and humanoid figure are simple forms of humanoid robots (HR) which, for instance, are already used in schools in Thailand. Further to being able to recite, answer and ask questions they have the ability to simulate emotions. With robots becoming more and more human-like, positive effects on learning can be observed. This phenomenon is called humanization and will be addressed in detail in Chapter 3.3.

Overall in this paper we will try to give an overview over the field of humanoid robots in education by looking at potential use cases (3.1), roles they can take on (3.2), advantages of humanization (3.3), and possible obstacles and drawbacks (3.4) respectively.

2 Methodology

We used the search strings "ai AND education AND humanoid AND robots" and "artificial AND intelligence AND education AND humanoid AND robots" considering the titles, abstracts and keywords of articles using the online databases Scopus and Web of Science. This yielded 92 results. Papers that mention education and humanoid robots independent of each other were excluded, since they were not concerned with the use of humanoid robots as a tool for education. Duplicates were also removed, as well as review articles, because the goal of this study is to investigate the results and arguments of applied research in this area. We reviewed seven articles with regards to their use cases, the role of the robot, advantages of humanization that are brought forward and the obstacles and drawbacks authors presented.

3 Results

In this section, contributions of the individual papers are assessed by examining the humanoid robots described from different angles. The section use cases (3.1) is concerned with the subject the robot teaches. Role of the robot (3.2) gives an overview of the social roles the robot can impersonate. In Advantages of humanization (3.3) the mechanisms of enhancing the learning experience through humanization are explored. Lastly, obstacles and drawbacks (3.4) goes into detail about the current limitations and shines a light on potential harmful consequences of this technology.

3.1 Use cases

Some research indicates that the application of humanoid robots in education is especially effective for a small range of subjects. Budiharto et al. (2017) claim that humanoid robots are a particularly good tool for education in the field of sciences, technology and math. Furthermore, it might be interesting to see how the requirements for robots vary across different levels of education.

The smart toy Maya is designed for preschool children and therefore does not teach a specific subject but a variety of different concepts that are age appropriate (Akdeniz & Özdinç 2021). When interacting with the toy it is possible to choose between four different fields of learning. The final version of the toy can help children deepen their knowledge about numbers, shapes, colors and animal names. One advantage of using AI and other smart technology is that the toy is able to identify areas in which a child needs more experience and therefore can provide individualized exercises.

Other authors also see humanoid robots as potential peer learners for children (van den Berghe et al. 2021). In an experiment conducted in the Netherlands, children at the age of five learned English. The robot, while it guided the child through the tasks, continuously repeated the newly learned words in a context of small games. For example, it says that the child needs to put ”three” (new English word) giraffes in the cage (in a game on a tablet).

In another study the main purpose of a humanoid robot was to teach and reinforce basic concepts of math to children in elementary school (Budiharto et al. 2017). Either the robot formulated short stories concerning transactions of fruits between multiple characters and asked the child for the number of fruits a specific character has in the end, or the child formulated stories and asked the robot. Apart from this, the robot was also able to engage in entertaining tasks such as storytelling and singing. This is an important part of education in elementary schools and therefore an essential use case.

Matsuura & Ishimura (2017) focus on higher education and use a humanoid robot to assist in lectures. A survey among science education students revealed that the majority of them can learn better in a more relaxed atmosphere. It also showed that many participants have difficulties following the long monologue by the lecturer and some learn better intuitively. Following these results they introduced a humanoid NAO robot that engages in a humorous dialogue with the lecturer to bring about a more friendly atmosphere and simultaneously explain the subject using a more familiar association or analogy that allows for understanding the topic better in an intuitive way.

Humanoid robots could also be used to support learning for autistic children since they are often overlooked in the school system, as Alcorn et al. (2019) investigate in their paper. Humanoid robots can be a good tool to aid the higher level of support autistic children need, without having to employ more teachers. Autistic people are also often interested in robots due to them being interactive but still predictable. They might therefore constitute a more comfortable environment for autistic learners to train social interactions in, that could then possibly be transferred to social interactions with humans.

Properly trained robots should be able to understand and react to the individual needs of students in order to allow for the best education and learning experience possible. For multilingual countries, like e.g. India, the robot would therefore need to speak multiple languages as well. As a result the robot could be used to translate between languages and therefore assist language learning (Barik et al. 2019).

Although the subject that is taught entirely or in part by the humanoid robots is not always clear in the reviewed papers, it is still clear that the fields of applications have a great range, from mathematics, over language learning to social skills. Besides, it could be seen how the procedure of learning varied from being playful for preschool children to having a dialogue between the robot and the lecturer in higher education.

3.2 Role of the robot

Humanoid robots can take on different roles in educational settings, such as acting as a teacher, tutor or peer. However, it is not always possible to put a precise label on it, as we will see in the following.

In a work done by Budiharto et al. (2017) a humanoid robot teacher that focuses on speech recognition was tested in elementary schools. The idea behind this robot is to imitate the capabilities of human teachers especially when handling question answering in noisy environments. Although the long term goal of this research is to create a robot teacher, the prototype used in the experiment has proven not to be able to fulfill that goal. Furthermore, the authors emphasize that robots used in education should be Socially Assistive Robots which usually occupy the role of a coach or teacher.

In Matsuura & Ishimura (2017), a humanoid robot is used as an assistant teacher in higher education. Engaging with the lecturer in Nazokake, a Japanese form of humorous riddle, the robot introduces an analogue topic to the content explained by the teacher. These analogue topics are supposed to be more familiar to the students and therefore meant to be understood more easily. Building this humorous atmosphere and including intuitiveness in the explanations improved the learning capacities of the students.

Such robot teachers could also be a great addition to the education systems of rather poor countries like India. Especially in rural areas, underdevelopment and low educational standards are a big issue, causing highly limited opportunities for people to learn even basic skills such as doing calculus or reading. The overall student-teacher ratio is highly uneven. It is therefore not uncommon to have up to 100 students taught by a single teacher and in many areas there is a high amount of single-teacher schools. Further knowledge and teaching methods are outdated and teaching materials available are insufficient. Smart robots, combining Machine Learning and Artificial Intelligence, can offer an important contribution to improve the education standards of countries like India. In contrast to humans, robots are able to work without pauses, need no further salary beyond investment and maintenance costs, have a high working memory enabling them to do several tasks at once and can be updated easily (Barik et al. 2019).

Another possible way to use humanoid robots is in learning with autistic children (Alcorn et al. 2019). Educators working with children on the autism spectrum stated that they could envision using humanoid robots as tutors or social interaction partners. However, it was important to them to have clear roles for these humanoid robots. The robot should not be seen as a toy. In order to limit the danger of drifting to a toy-like perception the devices should be used with clearly defined and evaluated goals. Furthermore, as children can have different skills that need to be worked on, the robots should be programmable to react suitably to different learners and work with them one on one. It was also made clear that the robots could not work alone with the autistic children as a human teacher would be needed to have supervision that everything is working fine. This would put the robots in the role of a tutor. Lastly, the educators also see potential in humanoid robots as social interaction partners for the autistic learners, as robots could be easier to interact with compared to real humans, but still help them to learn how their actions may affect others.

Akdeniz & Özdinç (2021) introduce Maya, a humanoid robot developed as a toy to support the learning of preschool children. The authors of the paper do not explicitly state what role they intended for their toy in relation to the children. However, they state their goal as supporting teachers with the toy rather than replacing them. Additionally, the toy has a similar height as the children it is made for, which might indicate that it is designed as a peer. The authors even suggest the toy might "evoke a sense of friendship" (Akdeniz & Özdinç 2021: p. 3). Because the authors set out to create a tutoring system and the authority of judging an exercise lies clearly on the side of the toy it can also be considered a tutor.

Some papers like Okanda et al. (2018) do not specify which role the robot is supposed to have. In their study they presented videos featuring a humanoid robot to four to five year old's and tested whether their answers change depending on the robots previous social appearance. In the main trial the robot showed the children objects and asked them questions about those, making it hard to assign the robot one of the common roles. It was neither a teacher or tutor, nor was it a peer or a friend.

In van den Berghe et al. (2021), a one-on-one experiment is conducted, where a humanoid robot learns a new language with a child. The robot was supposed to act as a slightly more knowledgeable peer, to be able to provide the children with feedback and support whenever it was necessary. Reasoning that many previous papers show the positive effects of peer learning in general the authors hoped that, by presenting the robot as a peer, the robot might be more socially present and therefore able to motivate the children more, which should result in better vocabulary learning results. The authors suggest that students might benefit when learning with the robot as a peer learner or tutor. The results seem to depend on the level of anthropomorphism the participant expresses, meaning how many human attributes a person assigns to a robot. The term of anthropomorphism will, however, be explained in more detail in the following part.

It becomes evident that many role classifications are rather fluid between teachers, tutors or peers. This is also due to the fact that some papers are still about hypothetical use cases without them being fully fleshed out yet. When designing a robot to be a peer it is its humanoid appearance that makes it most appealing, especially to younger children (reasons for this are given in 3.3 Advantages of humanization). On the other hand, the focus for robot teachers shifts more to their technical capabilities of easing the workload of human teachers. What also stands out is that humanoid robots in education are not yet expected to be used without supervision from a human educator, more on this in 3.4 Obstacles and drawbacks.

3.3 Advantages of humanization

In the process of humanization of an object or entity people ascribe human characteristics to it which therefore puts it on a similar level as humans. When this humanization happens in regards to robots, one often also talks about anthropomorphism.

As already mentioned in the previous part 3.2 Role of the robot, anthropomorphism plays a key role in the interaction with humanoid robots. When interacting with social robots people practice anthropomorphism to a certain degree. To put it more precisely: when interacting with a robot people tend to assign the robot, with whom they interact, human attributes in form of characteristics and behaviour. The phenomenon of anthropomorphism is neither new nor restricted to robots, humans have a tendency to ascribe human attitudes to toys, animals and many other entities. Since everyone practices anthropomorphism there is always an individual component involved. Depending on previous experiences, the attitude towards robots and own feelings or empathy, the degree of anthropomorphism can vary between individuals. Due to the tendency of people to show more positive reactions and attitudes towards robots if those show a higher resemblance to humans, it is possible to use the phenomenon of anthropomorphism beneficially in the human-robot interaction. This advantage is supported by the results of the paper by van den Berghe et al. (2021). Here a weak but nevertheless existing correlation between the degree of anthropomorphism and a comprehension score was shown. The participants, five-year-old children, learned a new language using the assistant help of a robot. Before and after the experiment they had to answer questions regarding anthropomorphism, such as "Do you think [...] the robot can feel pain?" or "Do you think [...] the robot understands when you say something?" (van den Berghe et al. 2021: p. 403). The larger the change in the degree of anthropomorphism was, meaning the more the children perceived the robot as a human, the better the results of the language test became.

Another part of the children's relation and perception towards the robot should be expected to be expressed through language. In the study Okanda et al. (2018) Children of a similar age, between four and five, were separated into two groups, from which one was shown a robot reacting very actively towards a human and their communication, while the other group saw a robot being unresponsive towards the human actions and communication attempts. However both groups reacted, talked and answered in a similar way when it was their turn to answer yes-no-questions from the robot without showing any differences in their way of communicating with a robot depending on their previous representation.

As mentioned previously, positive reactions of humans towards a robot increase the more the machine resembles a human, which might be the reason why the authors of the paper Akdeniz & Özdinç (2021) changed the appearance of the robot. Earlier versions of Maya had a non-human appearance but became, through the feedback of teachers, humanoid in order to attract the attention of children and allow for a more personal toy-child relationship. Other research has shown that the attraction of attention for humanoid toys is highest, if the toy is larger than other toys but smaller than the child (Tanaka & Kimura 2010). Considering these findings Akdeniz & Özdinç (2021) decided to make the robot a child-sized humanoid in the upper part of the body while lacking the lower part of the body completely. However, this decision was criticized by some of the teachers supporting this study. Even though a human's appearance might elicit a better response in humans, and especially children, it still has to be discussed what other effects their humanoid appearance might have. Possible risks and difficulties will be discussed in the last section of this paper.

However, it is possible that for certain minorities the application of humanoid robots is more reasonable than for other groups. Educators of autistic children indicated that the use of humanoid robots for learning could have some major advantages for these children. Autistic people often have difficulties in social interactions, since people can react in very different ways and this unpredictability of behaviour can be difficult to deal with for a person on the autistic spectrum. Unlike humans that behave in all kinds of ways, robots are predictable in their responses and always behave the same, while still being able to interact or respond in a human-like way. This makes them great to practice social interactions in a safer space before applying it to interactions with humans. Robots would therefore offer a safe form of interaction, which might motivate children to learn as they may not be afraid of making mistakes as with a human teacher (Alcorn et al. 2019).

In these cases a high level of anthropomorphism will be an advantage. The more human characteristics the robot would fulfill, the more secure the children could feel in their practice.

One can see that anthropomorphism is an important topic in the usage of humanoid robots. They try to reach learning advantages and maybe even social security for certain groups, acting like a human while being a robot. However, it is still in discussion which amount of human features is appropriate and whether too human-like robots might evoke new problems.

3.4 Obstacles and drawbacks

Although the majority of the reviewed papers spoke positively about the application of robots in education, the authors of some of the papers are cautious because of the limitations of the current technology and some of the possible effects of its application.

Despite offering many advantages for the Indian educational system it could still be difficult to find acceptance for a broad deployment of humanoid robots as people could be skeptical against robots teaching children and could also fear a decrease in job availability (Barik et al. 2019).

Using humanoid robots in a lecture setting has the goal of relaxing the atmosphere and improving the students' ability to learn (Matsuura & Ishimura 2017). Nevertheless, the authors voice the concern that robots can become a distraction from the lecture content and therefore have the opposite effect.

A concern voiced by van den Berghe et al. (2021) is that if the robot is humanized enough for children to form a relationship, this might happen at the expense of relationships with people. Although the technology currently in use has some obvious flaws and is still recognizable as a robot, the robot is already strongly humanized.

In the setting of using humanoid robots with autistic learners, one drawback could be that they might form an obsession with the robot and spend too much time with it (Alcorn et al. 2019). Educators of autistic children were especially worried that they would rather spend time with the robot and as a consequence neglect other soc ial interactions e.g. with their family which would be way more valuable to them. They also saw humanoid robots as a middle step between human-human interactions. Before they can be used, the children need to be prepared for their arrival by human teachers as many are anxious towards new people and activities. By learning social interactions with the predictable robot, they might not be well prepared for the great variety of human reactions, which is one more reason why the interaction with humans should always be the goal when using humanoid robots in this way.

Budiharto et al. (2017) also talk about the risk of young children who work closely with robots might get emotionally attached to them in an unhealthy manner. Additionally, they advocate that the lack of a universal theory of learning leads to humans being better suited as educators then pre-programmed robots. In order to increase the effects of education, they claim, future research should focus on making interactions between humans and machines as natural as possible.

Multiple papers about humanized robots in education warn that humanization could lead to strong and even unhealthy relations between children and robots. Therefore, future research should investigate the effects that are at play and find ways to minimize this risk.

4 Conclusion

The small number of publications that fulfill the criteria we applied (see 2 Methodology) shows that the use of humanoid robots in the field of education is not widespread at the present. Nevertheless, the examined papers show great potential in this technology. Since the spectrum of subjects the authors picked for their models and prototypes is so broad, including math, science, languages and social skills, and the range of ages in the experiments span from preschool to university, it is likely that humanoid robots will, in some shape or form, find a place in future education. With the age of students who are taught or supported in their learning by humanoid robots, also the appearance and role of the robots varied. Humanization of robots opens not only the opportunity to replace human teachers in areas where educators are undermanned but brings other advantages. When supporting human teachers, anthropomorphic robots show a significantly better learning experience compared to other tutoring systems. Additionally, for children with a special need of care, like children with autism, robots sometimes might be easier to handle than humans. The strong emotional connection between learners and the robots has raised some concern and requires careful oversight. Because human learning is a complex and developing field, more research needs to be conducted before a broad application of humanoid robots is possible in the field of education.

References

Akdeniz, Mevlüde & Fatih Özdinç. 2021. Maya: an artificial intelligence based smart toy for pre-school children. *International Journal of Child-Computer Interaction* 29. DOI: https://doi.org/10.1016/j.ijcci.2021.100347. https://www.sciencedirect.com/science/article/pii/S221286892100060X.

Alcorn, Alyssa M., Eloise Ainger, Vicky Charisi, Stefania Mantinioti, Sunčica Petrović, Bob R. Schadenberg, Teresa Tavassoli & Elizabeth Pellicano. 2019. Educators' views on using humanoid robots with autistic learners in special education settings in England. *Frontiers in Robotics and AI* 6. 107. DOI: 10.3389/frobt.2019.00107. https://www.frontiersin.org/article/10.3389/frobt.2019.00107.

Barik, Amar, Saumyakanta Khatua, Roshan Kishor Nikam, Subhash Pal & Parthrajsinh Gohil. 2019. Scope of implementing humanoid robots at educational institutes in India's perspective. In *2019 5th international conference on computing, communication, control and automation (iccubea)*, 1–6. DOI: 10.1109/ICCUBEA47591.2019.9128886.

Budiharto, Widodo, Anggita Dian Cahyani, Pingkan C.B. Rumondor & Derwin Suhartono. 2017. Edurobot: intelligent humanoid robot with natural interaction for education and entertainment. *Procedia Computer Science* 116. Discovery and innovation of computer science technology in artificial intelligence era: The 2nd International Conference on Computer Science and Computational Intelligence (ICCSCI 2017), 564–570. DOI: 10.1016/j.procs.2017.10.064. https://www.sciencedirect.com/science/article/pii/S1877050917321142.

Cóndor-Herrera, Omar, Hugo Arias-Flores, Janio Jadán-Guerrero & Carlos Ramos-Galarza. 2021. Artificial intelligence and tomorrow's education. In Tareq Z. Ahram, Waldemar Karwowski & Jay Kalra (eds.), *Advances in artificial intelligence, software and systems engineering*, 184–189. Cham: Springer International Publishing. DOI: 10.1007/978-3-030-80624-8_23.

Matsuura, Shu & Riki Ishimura. 2017. Chatbot and dialogue demonstration with a humanoid robot in the lecture class. In Margherita Antona & Constantine Stephanidis (eds.), *Universal access in human–computer interaction. human and technological environments*, 233–246. Cham: Springer International Publishing.

Okanda, Mako, Yue Zhou, Takayuki Kanda, Hiroshi Ishiguro & Shoji Itakura. 2018. I hear your yes-no questions: Children's response tendencies to a humanoid robot. *Infant and Child Development* 27(3). e2079 ICD-17-0041.R3, e2079. DOI: 10.1002/icd.2079. https://onlinelibrary.wiley.com/doi/abs/10.1002/icd.2079.

Tanaka, Fumihide & Takeshi Kimura. 2010. Care-receiving robot as a tool of teachers in child education. *Interaction Studies* 11(2). 263.

van den Berghe, Rianne, Mirjam de Haas, Ora Oudgenoeg-Paz, Emiel Krahmer, Josje Verhagen, Paul Vogt, Bram Willemsen, Jan de Wit & Paul Leseman. 2021. A toy or a friend? Children's anthropomorphic beliefs about robots and how these relate to second-language word learning. *Journal of Computer Assisted Learning* 37(2). 396–410. DOI: 10.1111/jcal.12497. https://onlinelibrary.wiley.com/doi/abs/10.1111/jcal.12497.

Chapter 15

AI-based virtual reality systems for helping learners with special needs

Yannik Heß, Michael Rau

People with special needs often face difficulties in daily life. This can be especially true for education as the system hardly accommodates for those needs. As our literature review shows, AI- and VR-based systems can help by providing scaffolds that support people with ASD to learn social skills and complex tasks such as driving, expose deaf and hard-of-hearing babies to language to prevent development delays and act as virtual sign language interpreters to alleviate restrictions hearing-impaired face with regards to their autonomy. We also realized that there is still research left to do for other special needs as they for example come with ADHD.

Keywords: Scaffolding, Virtual Reality, Virtual Agents, Special Needs

1 Introduction

Often, we say that education is empowering people, allowing us to be able to make up their own opinions as well as letting us choose our future and become "successful" in life. Therefore, it holds even more importance to provide inclusive education in our society.

> "Inclusion entails the right of all learners to quality education and the development of their full potential, regardless of special educational needs, disability, sex, social or economic backgrounds." (German Commission for UNESCO 2021)

In this article, we will focus exclusively on reporting the aspect of "special educational needs" (of course without denying the importance of all other aspects).

We will look at situations where there are struggles with inclusion in education and how technology might come in to help. This literature review is meant to gauge recent developments in the use of VR scaffolding and AI algorithms to provide learning support for people with special needs.

In Section 2, we will provide the background for our topic, after which, in Section 3, we explain our methodology. Section 4 presents exemplary VR systems. In the end, Section 5 will finish with some final thoughts and draw some conclusions for possible future research.

2 Background

2.1 Students with special (educational) needs

Some of the most prominent groups of people that have special needs in education are those being deaf or hard-of-hearing, having attention-deficit / hyperactivity disorder (ADHD), and being diagnosed with autism spectrum disorder (ASD). All their learning processes are to various degrees – but usually higher than those of students without special needs – dependent on environmental circumstances.

Roughly summarized, students with special needs face two major problems: They have problems accessing educational contents due to restricted abilities and they struggle with regard to social interactions and communication (especially with peers that do not have the same special needs). However, social interactions are an important part of learning as well.

2.1.1 ASD

In people with ASD, the environmental impact on learning is probably most noticeable. They are more likely to get distracted and overwhelmed by sensory information from the environment, thus making attention focusing on the content to be learnt much more difficult for students with ASD. Due to less social skill, students with ASD have a hard time participating in play with peers, which is an important source for learning as well. (Dahl et al. 2021)

Additionally, individuals with ASD seem to have different capacities in working memory compared with TD (typically developing) individuals. Individuals with ASD perform worse on tasks related to working memory, with an equally altered performance under different levels of cognitive load. (Zhang et al. 2017)

While ASD is one of the most studied intellectual disabilities today, its diagnosis and treatment are still challenging due to the large spectrum of forms ASD can come in. That is why standardized and generalized tests are sometimes criticized for not doing justice to the diversity of ASD. Additionally, those tests are only applicable for three-year-olds or older children. Here, technology might come in handy. It can be used to analyze non-verbal communication and other specific inputs like eye movements to detect repeated behavioral patterns stereotypical for people with ASD. (Guerrero-Vásquez et al. 2017)

On the treatment side, VR and AI technology can be individually tailored to the needs of people with ASD to both comfort them to feel safe and decrease distractions but also challenge them at a suitable level to allow learning of especially social skills. Technology may also act as a low-level starting point to mediate interaction with peers. (Dahl et al. 2021)

2.1.2 Deaf and hard-of-hearing

The COVID-19 pandemic raised special awareness to a lot of different problems in our society and not the least in our educational system. For deaf and hard-of-hearing students, the hurdles were strongly emphasized (c.f., Schafer et al. 2021). However, Schick et al. (2005) already pointed out how deaf and hard-of-hearing students are "left behind" in 2005. Often, deaf and hard-of-hearing students are highly dependent on educational interpreters to not only access the educational contents but the distributed social communication to and from (only) hearing peers as well. Due to this, learning is generally more difficult for deaf and hard-of-hearing students. There are delays in translation and the need to split visual attention between interpreter and visual material. Furthermore, interpreters need to be both knowledgeable in taught contents and at the same time convey other information from teachers and hearing students that is not in what they say but how they say it. (Schick et al. 2005)

VAs could interpret specific educational content or mediate between deaf and hard-of-hearing students and their non-signing peers to increase the deaf and hard-of-hearing's access and independence in schools.

Moreover, the fact that 91.7% of deaf children are born into non-signing families comes with additional problems too. For example, the lack of exposure of young babies to language might require early intervention to support fundamental language learning stages. Technology might help by providing signing avatars that can expose young babies to (sign) language (Nasihati Gilani et al. 2019).

2.1.3 ADHD & ADD

There is a relatively high known prevalence of children with attention-deficit / hyperactivity disorder (ADHD) of 8.5% to 9.5% – compared to, for example, 1.1%

to 2.5% of children with ASD – in the United States (Zablotsky et al. 2019). ADHD is divided up into two main types: the hyperactive-impulsive one and the inattentive one (formerly known as ADD). Both types can also be present in a combined form of ADHD. People with ADHD can show a wide variety of different features, some of the ones that are especially interesting regarding education include performance inconsistency, poor working memory, poor incentival motivation and learning difficulties as well as poor concentration, task impersistence, disorganization and forgetfulness for the inattentive type or implusivity and overactivity for the hyperactive-impulsive type. (Selikowitz 2021)

As a study by Sarver et al. (2015) showed, the stereotypical fidgety behavior of ADHD children is only apparent while using executive functions of the brain. While the widespread believe is that they are not interested and do not concentrate when learning, the movements actually help them to retain a certain level of alertness (Sarver et al. 2015). Therefore, creating learning units with compensatory movements already in mind could be of great advantage for learners with ADHD. For certain topics VR learning enviroments could be the right tool.

2.2 Learning: motivation & scaffolding

For learning, motivation and engagement are essential. Traditionally, we distinguish between extrinsic and intrinsic motivation to specify its source. Although we often only talk about these as mere binary options, the source of motivation can be seen as rather lying on a spectrum somewhere between extrinsic and intrinsic. (Dahl et al. 2021)

In general, Dahl et al. (2021) propose that a balance between extrinsic and intrinsic motivation is desirable. Too much extrinsic motivation can impede the acquisition of self-reliant motivation. However, intrinsic motivation is very difficult to foster and learning by extrinsic motivation can help to spark interest in topics where intrinsic motivation can then be developed. (Dahl et al. 2021)

In addition, feelings of achievement are central aspects of learning which require a certain degree of autonomy to allow students to get the impression of self-earned results (Dahl et al. 2021). The fact that play provides some sort of freedom, might be a reason why it can be quite useful in educational contexts.

Gamification – due to the introduction of external rewards –, Dahl et al. (2021) mentioned, is noticeably more on the extrinsic side of the motivational spectrum. Additionally, it uses mostly behaviorism as theoretical background by focusing on the student's behavior to certain stimuli rather than cognitive processes (Dahl et al. 2021). While such concepts find application in VR systems for educational contexts as well, we would like to focus on cognitive approaches to the use of VR and VAs (virtual agents) and how they can function as scaffolding to facilitate situations and environments where engagement and motivation may be precisely reinforced.

The perspective of scaffolding on cognitive processes such as learning can be understood quite literally as considering them against the backdrop of the structures – resources, knowledge, and experiences – that an individual or a group of individuals has at their disposal. A behaviorist approach focuses on stimuli input and behavioral outputs only, whereas we think it is important to assess and create technological systems – that shall help students with special needs – with internal cognitive processes as well as the environment's significance in mind.

As Dahl et al. (2021) point out, a merely behaviorist approach has the potential to reduce the subjects in question to something that can be "conditioned and manipulated into behaving according to ideals they themselves do not identify with" (Dahl et al. 2021). We, therefore, like to take on viewpoints from 4E cognition and scaffolding as well to gain a more complete overview.

3 Methodology

To conduct our literature search, we employed the use of the websites *Scopus* and *Web of Science* with the search terms listed in Table 1. According to the multiple steps depicted in Figure 1, we filtered the 95 resulting papers by a publishing dates that were not older than 2016, removed duplicates, filtered for accessibility and excluded entire conference proceedings (where only all the papers combined included the search terms), non-English papers and filtered for topic relevancy. We screened the 30 papers left fully and filtered them according to the topical guidelines of the book and our topic (excluding: not AI, not education, augmented reality instead of other forms of VR). We filtered out augmented reality (AR) because we deemed it more efficient to concentrate only on selected forms of VR. AR – in bringing the virtual world into the real world – would provide many more aspects to consider, which would extend the scope of this paper. Finally, we added an additional paper by hand since it provided technical background information to another paper, so that we reached 10 suitable papers.

4 Results: VR scaffolds for people with special needs

In this section, we take a look at a few examples for VR-based scaffolding that serve as learning support for people with special needs. These systems utilize

Table 1: Topics and corresponding search terms

Topics	Search terms
AI	”ai” OR ”artificial intelligence” OR ”intelligent agent”
AND	
VR & VA	”vr” OR ”virtual reality” OR ”virtual agent”
AND	
Education	”education” OR ”learning”
AND	
Special Needs	”ASD” OR ”autism” OR ”sign language” OR ”deaf” OR ”hard-of-hearing” OR ”special needs” OR ”ADHD” OR ”attention deficit hyperactivity disorder” OR ”attention deficit disorder” OR ”attention deficit syndrome”

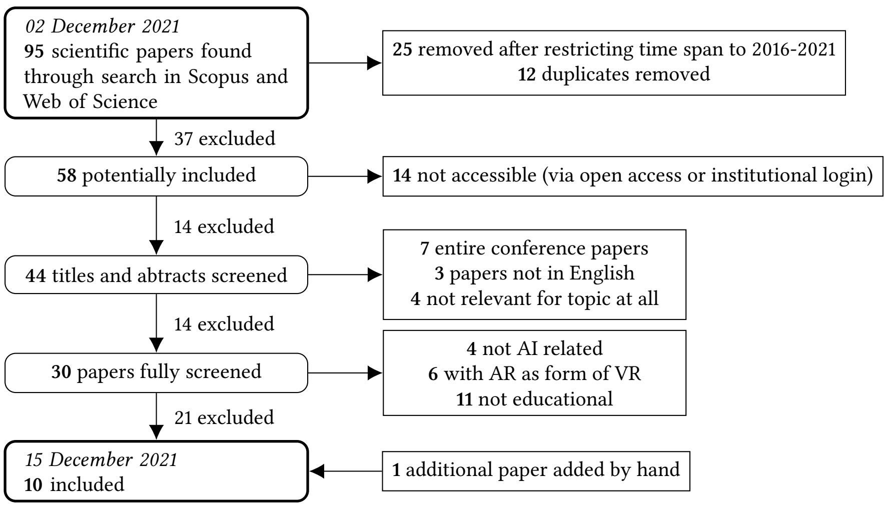


Figure 1: Filtering process of literature for reviewing (modified PRISMA diagram inspired by Zawacki-Richter et al. 2019)

AI algorithms as a way to efficiently support learners according to individual requirements.

4.1 VR-based driving simulators for individuals with ASD

One of the challenges faced by individuals with ASD is driving – which is often a necessitiy for an autonomous life with minimal support – since it requires a sharp focus and attention due to the amount of input signals involved.

Teens with ASD have been found to make more driving errors than their neuronormative peers (Zhang et al. 2017). Significant differences were also found regarding unsafe gaze patterns and higher anxiety, resulting in lower attention. Furthermore, they were inclined to respond slower during steering, identify fewer social hazards and show a lower multi-tasking ability. (Fan et al. 2018)

Multiple studies have been conducted to evaluate possible learning support approaches through the use of technology. We encountered three papers that focus on VR-based approaches featuring immersive driving simulators. For example, Zhang et al. (2017) intended to analyze the (notoriously low-capacity) cognitive workload of participants with ASD in their study to build a framework for future assessment.

To accurately gauge cognitive load, Zhang et al. (2017) employed a multimodal approach to collect data. They found that the combination of different modalities, such as an EEG (electroencephalogram), eye gaze analysis, physiological information (heart rate) and performance evaluation together with a subjective evaluation by the participants achieved the most promising results to measure cognitive load during and after the driving test. A machine learning (ML) algorithm that first captured and evaluated eye gaze data, to secondly evaluated performance and physiological information and thirdly analyzed EEG data proved to be the most accurate of all tested ML-algorithms.

In a different study, Fan et al. (2018) concentrated on classifying the emotional state of participants by analyzing workload and affective state EEG data through a k-nearest neighbor algorithm. The researchers found that the proposed model based on EEG activation is able to detect with high accuracy subjects' states of low engagement, low enjoyment, high frustration, and high workload, while boredom seemed to be harder to measure precisely. Nonetheless, this approach is a promising path to analyzing the affective state of learners. (Fan et al. 2018)

With an accurate analysis of cognitive workload, a further algorithm could be envisioned to modify tasks on the fly by introducing more input or reducing it. Another promising approach is the model of an algorithm for VR-based driving simulators proposed by Bian et al. (2016). They proposed a system that would utilize an automatic analysis of performance data in order to gauge a participant's engagement level with tasks. Depending on the analysis of the participants' affective state, the algorithm would then adjust the task's difficulty. These proposed algorithms serve to achieve a more effective learning experience for the learner and thus remove barriers for people with ASD when it comes to acquiring driving skills. (Bian et al. 2016)

4.2 Social skills training for autistic children with the help of VAs

Other examples for AI-assisted VR systems are AVATAR and ECHOES. Both use VAs as agents that directly interact with autistic children in therapeutic and classroom settings respectively. The VAs are displayed on screens in front of the children and the interaction with the VAs are held very simple while immersive VR headsets are not used due to the young developmental age (4-7 in the case of ECHOES) of the systems' users. Additionally, "the environment of the child is not invaded at any time" Guerrero-Vásquez et al. (2017), which allows the child even more control over the situation. However, in other situations VR headsets might be advantageous due to being more effective in blocking out distracting factors (Porayska-Pomsta et al. 2018, Guerrero-Vásquez et al. 2017).

On first sight, both systems are very similar – the basic idea being that the VAs are acting as responsive social partners. They are specifically designed to fit the needs of children with ASD and are personally adapted to each child. Though, when going more into details, there are a few key aspects distinguishing both systems. (Porayska-Pomsta et al. 2018, Guerrero-Vásquez et al. 2017)

AVATAR combines AI with human intelligence: An expert system mediates between a therapist – monitoring the child's actions – and the interface of the child. The expert system also promotes certain interactions and assesses the child's progress. Furthermore, AVATAR is first and foremost built for interactions where the child imitates the actions of the virtual avatar in the real world. (Guerrero-Vásquez et al. 2017)

In ECHOES, however, the child will do tasks that involve imitation in the virtual world of the avatar itself. This introduces, for example, turn-taking into the interaction. Moreover, a teacher (or researcher in the study) will be present during the use of ECHOES allowing turn-taking with and acknowledgements from humans to be part of the interaction as well. After the teacher chooses the tasks, Andy – the VA of ECHOES – will act independently, contrary to AVATAR where the therapist can influence the VAs behavior and appearance at any time. Andy can proact – motivating the child and sustaining their attention – and react – accommodating for specific needs and changing the behavior of the child. (Porayska-Pomsta et al. 2018)

Besides introducing the ECHOES system, Porayska-Pomsta et al. (2018) also examined how the use of ECHOES influenced the children's social behavior. They found increased initiations of and responses to bids for social interaction with Andy and the present adult. The increased interaction with the latter also prevailed afterwards as independent tasks before and after the use of ECHOES revealed. Despite the increase not being statistically significant due to the very diverse subject group, real world transfer of increased social abilities were reported for some subjects. This shows promising potential for systems like ECHOES to help children with ASD to train their social skills. (Porayska-Pomsta et al. 2018)

4.3 VR systems for hearing impaired children in educational contexts

For examining AI-based VR systems and VAs helping deaf and hard-of-hearing in learning situations, we found two papers in our search for literature and added an additional one that elaborates on the technical aspects of one of the two found.

Nasihati Gilani et al. (2019) researched how a signing VA might ”Engage a Baby’s Attention” (Nasihati Gilani et al. 2019) to investigate whether such an avatar can be used for interventions to expose young deaf and hard-of-hearing babies (6-12 months) to language. This is important as they claim because, at this age, the brain (both of deaf/hard-of-hearing and hearing babies) has a special sensitivity for special rhythmic temporal patterns indispensable for later language development. Children only exposed to language later on will most likely have developmental delays that are hard to gain on. (Nasihati Gilani et al. 2019)

Their so-called RAVE system uses both a physical robot and a signing humanoid avatar on a monitor. The described paper focuses on the latter, and to which extent it might be able to engage the baby’s attention and stimulate behavioral responses – linguistic, social/gestural, and sustained visual attention. Nasihati Gilani et al. (2019) also equipped the RAVE system with perceptual modules that allow the system to adapt its behavior in real-time. In particular, the baby’s state of engagement was measured via eye-tracker, thermal infrared imaging and a human observer interface. A dialogue manager receives the sensory signals and provides a plan for the two agents (robot and avatar) to execute to attain a certain goal (e.g., baby looking at the avatar). If the baby changes its behavior the dialogue manager can update the plan. This shall ensure that the interaction between baby and avatar (and robot) is socially contingent. The dialogue manager uses a rule based policy to find the right sequence of reactions in a state space for the 460 possible combinations of input signals. (Nasihati Gilani et al. 2018)

Nasihati Gilani et al. (2019) found that the RAVE system (specifically the avatar) can engage the babies’ attention and that the mode in which it was reciting nursery rhymes – which were ”built in the specific rhythmic temporal patterning unique to phonetic-syllabic units in natural language phonology” (Nasihati Gilani et al. 2019) – resulted in the most sustained visual attention and (proto-)signs compared to other modes. Additionally, parents’ involvement seemed to further the system’s impact on language learning. (Nasihati Gilani et al. 2019)

In the second paper we found, Wen et al. (2021) introduce a system they created that uses a so-called ”triboelectric smart-glove” and AI-based analytics to enable bidirectional communication between signers and non-signers. The glove uses deliberately placed sensory fields from which's activation patterns a CNN (convolutional neural network) can detect learnt sentences and words in ASL (American sign language). The CNN can accurately identify 50 words and 20 sentences of which a significant amount is too similar to be identified through similarity-based methods. As the CNN of Wen et al. (2021) is segmentation-assisted, never-seen sentences – with the same words, but in different orders – can be detected as well. VR can be used as an interface for the interaction to enhance the communication and make it more immersive (Wen et al. 2021).

5 Conclusion

In our paper we aimed at examining AI- and VR-based systems as educational supports for people with special needs and focused on scaffolding concepts. In general, people with special needs can require very diverse kinds of support. This is why the use of AI is important to deliver personally tailored content to maximize the potential learning effects and/or help in learning situations.

Analysis of cognitive load during VR-based driving tasks can help to adjust task difficulty and improve engagement for individuals with ASD. Furthermore, people with ASD can benefit from VR-based training to improve social skills.

Deaf and hard-of-hearing students could gain more independence and access by having a virtual sign language interpreter that is able to mediate between signing and non-signing peers. For young deaf or hard-of-hearing babies, intervention in early language stages could be facilitated by the use of virtual signing avatars (like part of the RAVE system) to compensate deficits in language exposure often found within non-signing families.

Quite surprisingly, we weren't able to find any relevant papers in the scope of this article with regards to ADHD while using similar strategies as for deafness and hard-of-hearing as well as ASD. This came especially unexpected because ADHD is more than three times as likely to occur than ASD. Therefore, we suggest looking further into possible applications of AI-assisted VR systems helping to alleviate learning difficulties. Some of the findings for ASD might also be applicable for people with ADHD as both overlap in some of the symptoms and children are frequently diagnosed with both (Zablotsky et al. 2020).

We conclude that scaffolding in the form of AI-guided VR systems is a promising approach to not only enhance the education of people with special needs, but can also serve to improve accessibility to education and a greater independence. Technology like this has been proven to aid people with ASD and hard-of-hearing students. Additionally, there might be other groups with special educational needs out there, that we have not considered yet, that could greatly benefit from educational support through AI and VR.

References

Bian, Dayi, Joshua Wade, Zachary Warren & Nilanjan Sarkar. 2016. Online engagement detection and task adaptation in a virtual reality based driving simulator for autism intervention. In Margherita Antona & Constantine Stephanidis (eds.), *Universal access in human-computer interaction. users and context diversity*, 538–547. Cham, Switzerland: Springer International Publishing. https://link.springer.com/chapter/10.1007/978-3-319-40238-3_51.

Dahl, Christine, Henrik Skaug Sætra & Anders Nordahl-Hansen. 2021. Computer-aided games-based learning for children with autism. In Xiaowen Fang (ed.), *HCI in games: experience design and game mechanics*, 145–158. Cham, Switzerland: Springer International Publishing. https://link.springer.com/chapter/10.1007%2F978-3-030-77277-2_12.

Fan, Jing, Joshua W. Wade, Alexandra P. Key, Zachary E. Warren & Nilanjan Sarkar. 2018. EEG-based affect and workload recognition in a virtual driving environment for ASD intervention. *IEEE Transactions on Biomedical Engineering* 65(1). 43–51. DOI: 10.1109/TBME.2017.2693157.

German Commission for UNESCO. 2021. *Inclusive education*. https://www.unesco.de/en/education/inclusive-education (7 January, 2022).

Guerrero-Vásquez, Luis Fernando., Jack F. Bravo-Torres & Martín López-Nores. 2017. Avatar ”autism: virtual agents to augment relationships in children”. In *2017 IEEE xxiv international conference on electronics, electrical engineering and computing (INTERCON)*, 1–4. DOI: 10.1109/INTERCON.2017.8079705.

Nasihati Gilani, Setareh, David Traum, Arcangelo Merla, Eugenia Hee, Zoey Walker, Barbara Manini, Grady Gallagher & Laura-Ann Petitto. 2018. Multimodal dialogue management for multiparty interaction with infants. In *Proceedings of the 20th ACM international conference on multimodal interaction* (ICMI ’18), 5–13. Boulder, CO, USA: Association for Computing Machinery. DOI: 10.1145/3242969.3243029. https://doi.org/10.1145/3242969.3243029.

Nasihati Gilani, Setareh, David Traum, Rachel Sortino, Grady Gallagher, Kailyn Aaron-Lozano, Cryss Padilla, Ari Shapiro, Jason Lamberton & Laura-Ann Petitto. 2019. Can a signing virtual human engage a baby's attention? In *Proceedings of the 19th ACM international conference on intelligent virtual agents* (IVA '19), 162–169. Paris, France: Association for Computing Machinery. DOI: 10.1145/3308532.3329463. https://doi.org/10.1145/3308532.3329463.

Porayska-Pomsta, Kaundefinedka, Alyssa M. Alcorn, Katerina Avramides, Sandra Beale, Sara Bernardini, Mary Ellen Foster, Christopher Frauenberger, Judith Good, Karen Guldberg, Wendy Keay-Bright, Lila Kossyvaki, Oliver Lemon, Marilena Mademtzi, Rachel Menzies, Helen Pain, Gnanathusharan Rajendran, Annalu Waller, Sam Wass & Tim J. Smith. 2018. Blending human and artificial intelligence to support autistic children's social communication skills. *ACM Transactions on Computer-Human Interaction* 25(6). DOI: 10.1145/3271484. https://doi.org/10.1145/3271484.

Sarver, Dustin E, Mark D Rapport, Michael J Kofler, Joseph S Raiker & Lauren M Friedman. 2015. Hyperactivity in attention-deficit/hyperactivity disorder (ADHD): impairing deficit or compensatory behavior? en. *Journal of abnormal child psychology* 43(7). 1219–1232.

Schafer, Erin C., Andrea Dunn & Alexandra Lavi. 2021. Educational challenges during the pandemic for students who have hearing loss. *Language, Speech, and Hearing Services in Schools* 52(3). 889–898. DOI: 10.1044/2021_LSHSS-21-00027. https://pubs.asha.org/doi/abs/10.1044/2021_LSHSS-21-00027.

Schick, Brenda, Kevin Williams & Haggai Kupermintz. 2005. Look who's being left behind: educational interpreters and access to education for deaf and hard-of-hearing students. *The Journal of Deaf Studies and Deaf Education* 11(1). 3–20. DOI: 10.1093/deafed/enj007. https://doi.org/10.1093/deafed/enj007.

Selikowitz, Mark. 2021. *ADHD: the facts*. Oxford University Press.

Wen, Feng, Zixuan Zhang, Tianyiyi He & Chengkuo Lee. 2021. AI enabled sign language recognition and VR space bidirectional communication using triboelectric smart glove. *Nature Communications* 12(1). 53–78. DOI: 10.1038/s41467-021-25637-w. https://doi.org/10.1038/s41467-021-25637-w.

Zablotsky, Benjamin, Lindsey I. Black, Matthew J. Maenner, Laura A. Schieve, Melissa L. Danielson, Rebecca H. Bitsko, Stephen J. Blumberg, Michael D. Kogan & Coleen A. Boyle. 2019. Prevalence and trends of developmental disabilities among children in the united states: 2009–2017. *Pediatrics* 144(4). e20190811. DOI: 10.1542/peds.2019-0811. https://doi.org/10.1542/peds.2019-0811.

Zablotsky, Benjamin, Matthew D. Bramlett & Stephen J. Blumberg. 2020. The co-occurrence of autism spectrum disorder in children with ADHD. *Journal of At-*

tention Disorders 24(1). PMID: 28614965, 94–103. DOI: 10.1177/1087054717713638. https://doi.org/10.1177/1087054717713638.

Zawacki-Richter, Olaf, Victoria I. Marín, Melissa Bond & Franziska Gouverneur. 2019. Systematic review of research on artificial intelligence applications in higher education–where are the educators? *International Journal of Educational Technology in Higher Education* 16. 39. DOI: 10.1186/s41239-019-0171-0.

Zhang, Lian, Joshua Wade, Dayi Bian, Jing Fan, Amy Swanson, Amy Weitlauf, Zachary Warren & Nilanjan Sarkar. 2017. Cognitive load measurement in a virtual reality-based driving system for autism intervention. *IEEE Transactions on Affective Computing* 8(2). 176–189. DOI: 10.1109/TAFFC.2016.2582490.

Part IV

Addressing Special Needs

Chapter 16

AI approaches towards reducing the barriers for learners with sensory or physical impairments

Kamran Vatankhah-Barazandeh & Christian Meißner

One of the aspects of the recent surge in AI development is the application of AI technology for improvements in education. Among those, there has been research conducted which aims to increase the quality of education for learners with disabilities affecting their vision, hearing, or physical abilities. By showcasing openly accessible work found through systematic web searches, this review aims to provide an overview of achievements made throughout the past two decades in this area. It is divided into research improving the education for people with hearing impairments, vision impairments and physical impairments. The most significant improvements can be observed in the areas of image and speech recognition resulting in enhanced communication, promoting participation and improving the learning environment itself.

Keywords: vision, hearing, physical, impairment, disability

1 Introduction

In Sparkes (2019), it is reported that people with a disability are significantly less likely to hold a degree. This discrepancy is not only observed between the general populations of those with a disability and those without, but also specifically for sensory disabilities affecting hearing, those affecting vision, as well as physical disabilities.

It is important to note that every form of disability occurs on a spectrum and that we therefore often need individualized support systems that are tailor-made for the present set of impairments.

This review aims to provide an overview of the research on utilizing AI in Education for learners with disabilities, which might help close the aforementioned educational gap by reducing the obstacles teachers and learners have to work around. It includes articles on using AI to improve the learning environment for those with a disability either affecting one of their senses, in particular hearing and vision. We also take a look at AI approaches toward education for people with physical disabilities, but there is a much smaller amount of research in this area, since the focus currently seems to be on rehabilitation methods. The broad field of specific learning/language disabilities is not included in this review. We also do not consider any kind of AI driven diagnostic tools, since sensory and physical impairments are, in contrast to learning/language impairments, typically diagnosed by a physician within the first year of a person's life. Approaches for diagnosing learning/language impairments are also not included in this review.

With Artificial Intelligence being a wide field, we included articles that refer to any sub-field of Artificial Intelligence, leaving the definition to the respective authors. We left out vague mentions of AI that do not specify which particular method would be used. We do not focus on any part of education exclusively.

2 Methodology

The sources were collected from different databases using their respective search functions, namely Google Scholar, Scopus and Citeseer.

We used "AI Education" and "Artificial Intelligence Education" as search terms, each in combination with the terms listed in table 1.

Table 1: Search terms

hearing impairment
hearing disability
visual impairment
visual disability
physical impairment
physical disability
special needs

Our review only includes articles which were published in English and which can be accessed openly or for free with university credentials. However, we made an exception for blog entries which can be considered highly influential to the field. We also used data published by health organizations like the WHO. We only included articles published after 2000 to ensure some topicality.

3 Results

The presentation of our results is structured based on the definition used by Garg (2020). Accordingly, disabilities are categorized as hearing, visual and physical disabilities, as well as specific learning/language disabilities, the latter not being included in the search process for this review.

3.1 Hearing impairments

Depending on the severity of the impairment as well as the learning situation, there have been multiple approaches to improve communication between a hearing teacher and a learner with a hearing impairment. For some patients, modern hearing aids can improve hearing to a good level. This group of people grows steadily, as improvements in technology are translating into over-the-counter consumer products, and many expect AI methods to improve future versions even further (Zeng 2017). However, these solutions do not apply to every learner, as some cases of hearing loss are total or too severe for hearing aids to be a remedy.

For those who can not hear or understand their teacher in their educational environment, AI methods of transcribing or translating their speech can be helpful. Speech recognition tools use AI to allow real time translation from spoken into written language. One example of this is shown in Roach (2018). In this case, Microsoft Translator is used for real-time transcription of a biology lecture.

As mentioned in the paper, transcription is not always ideal, since the learner has to focus on the text output. This can lead to situations where they will have to divide their attention between what the teacher says and the lecture material. The paper mentions that being able to save the transcript can increase accessibility, if the lecture is also recorded. The learner could then rewind or pause in such situations.

Hearing aids as well as real-time translation tools are both designed to be integrated into a preexisting learning environment, like a classroom, without significant changes to the teaching process itself. There is also research being conducted on the use of e-learning platforms utilizing AI, which require more change to the structure of teaching the subject, but promise a better learning experience (Drigas et al. 2008) The systems in these examples are tailored towards those who speak Greek Sign Language, and intend to learn English. In these cases, AI (in the form of a knowledge-based system using fuzzy logic) is used to evaluate the learner. Systems like these can be seen as a subset of e-learning approaches.

3.2 Visual impairments

As pointed out in the introduction, visual impairments exist on a spectrum. Data published by the WHO ("Blindness and vision impairment" 2021) show that visual impairments are a rather popular occurrence with at least 2.2 billion people having to deal with some kind of visual impairment. Here, total blindness is a rare exception. However, it is important to point out that early onset visual impairments do correlate with developmental delays in motor functioning, cognitive areas and language skills. These delays can manifest in lifelong consequences such as an overall lower level of educational achievements. As shown in "Blindness and vision impairment" (2021), Mboshi (2018), visual impairments in adults are associated with higher rates of depression and anxiety as well as problems in social aspects and mobility, resulting in a severe impact on quality of life.

Many products targeting people with visual impairments aim at increasing the quality of life, which often comes with a positive impact on the educational potential. An example reported by Wu (2019) is Microsoft's App "Seeing AI", which is designed as a tool to improve visually impaired people's day-to-day life by reading out texts captured by the phone's camera, identifying money when paying in cash, recognizing faces and giving a picturesque description of one's surroundings. Even though this app is not made specifically for educational purposes, it can still be used, for example, to convert written homework assignments into audio form. Therefore, we observe a fluid boundary between tools supporting education and tools supporting everyday quality of life.

One possible approach towards the development of tools supporting the educational process of people with visual impairments is to build upon pre-existing technologies that have been in use for some time. In 2011, Robles Bykbaev et al. (2011), a research team from Ecuador, used the fairly common tactile writing system Braille to come up with novel ideas for educational support such as the "Braille Calculator". It is equipped with Braille buttons and a module to read out the calculated results. Additionally, the calculator is claimed to interact with a not further described educational system when connected to a computer.

Patra & Chander (2021) emphasize the importance of Optical Character Recognition (OCR) for a variety of different impairments. OCR describes the mechanical or electronic conversion process consisting of first scanning a presented image containing handwritten, typed, or printed text, followed by the recognition of the shapes of the displayed characters using AI technologies. The final outcome of the OCR process is a text file containing the recognized characters of the original image. OCR in combination with a text-to-speech system is of great help unlocking access to educational resources for people with visual disabilities. An example of a combined approach of OCR and text-to-speech is Microsoft's "Seeing AI" mentioned above.

Patra & Chander (2021) also point out the advantage of speech recognition systems in the form of virtual assistants to control and interact with devices. People with visual disabilities can simply ask their virtual assistant of choice to e.g. google an unfamiliar technical term via a voice command rather than via keyboard and computer screen.

Mboshi (2018) discusses educational provision of learners with visual impairment in inclusive classrooms. She underlines the importance of special software like "JAWS" that enables visually impaired students to experience graphical user interfaces such as a website seen on a PC screen. The visually displayed control elements and text is omitted to a non-visual output device. This translation can either take place via acoustic speech synthesis or via a braille display, a device presenting the information in a tactile manner.

Blind born children typically have problems developing communication skills as well as focusing on other people. As Grayson et al. (2020) report this results in up to two-thirds of blind born children being diagnosed within the autism spectrum. Morrison et al. (2021) relied on PeopleLens, an advanced version of Holo Lens (Grayson et al. 2020), which is a head-mounted Augmented Reality device housing cameras and sensors. It gathers spatial information and sends them to a server where the images are analyzed by perception algorithms. The processed information is then presented to the user via 3D spatialized audio. PeopleLens provides a mental map that allows the wearer to locate, track and identify people in a 4 meter radius around them. The world state model is updated in real time based on the user's field of view. It also provides the wearer with an analysis of people's body posture and gaze direction. It has been designed to help children socialize in school and ultimately enables them to use their head and body in a more convenient way to signal intentions and initiate interaction. This task-independent nature of PeopleLens allows for sustained improvements in blind children's everyday social life beyond the direct usage of the device.

3.3 Physical impairments

Approaches toward simplifying education for the physically disabled are less common than for the previously mentioned fields. This might be because they are mostly needed where physical motion is required, and because the solutions would have to be highly individual to accommodate the respective disability. Most research seems to be directed towards rehabilitation, and is not included in this review. For some patients, there is a communication barrier which is dealt with by using a haptic user interface which translates into language. This can be made easier with speech recognition (Huang et al. 2014, Patra & Chander 2021: 4), allowing easier communication.

4 Conclusion

As established in the *Universal Declaration of Human Rights* (1948), Art. 26, everyone has a right to education. To ensure this, it is important that inclusion is improved wherever it is possible. Beyond just allowing access for everyone, good education for those with a disability can be improved with new technology, provided research is possible and funding is provided. The field of AI in Education for the disabled covers many aspects, from improving communication to changes to the learning environment itself. The rapid development of new AI applications provides a wide range of opportunities for improvement, some of which have led to the achievements presented in this review, others still unexplored. Especially for physical disabilities, where highly individual solutions are often needed, there is still a comparatively low amount of research focusing on developmental support. For the hearing impaired, technology profits from rapid improvements in voice recognition and machine translation, which in turn often stem from research by large tech companies. The ubiquity of smartphones with camera and even computer vision capabilities, as well as modern sensory technology, also leads to faster deployment of new technology aiding the visually impaired. Again, much of this is made possible in the wake of well-funded research towards consumer electronics.

References

Blindness and vision impairment. 2021. *World Health Oranization*. Blindness%20and%20vision%20impairment (5 January, 2022).

Drigas, Athanasios, Dimitris Kouremenos & John Vrettaros. 2008. Teaching of English to Hearing Impaired Individuals Whose Mother Language Is the Sign Language. en. In Miltiadis D. Lytras, John M. Carroll, Ernesto Damiani & Robert D. Tennyson (eds.), *Emerging Technologies and Information Systems for the Knowledge Society*, vol. 5288, 263–270. ISSN: 0302-9743, 1611-3349 Series Title: Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg. DOI: 10.1007/978-3-540-87781-3_29. http://link.springer.com/10.1007/978-3-540-87781-3_29 (6 January, 2022).

Garg, Shalini. 2020. Impact of artificial intelligence in special need education to promote inclusive pedagogy. *International Journal of Information and Education Technology* 10. 523. DOI: 10.18178/ijiet.2020.10.7.1418.

Grayson, Martin, Anja Thieme, Rita Marques, Daniela Massiceti, Ed Cutrell & Cecily Morrison. 2020. A dynamic ai system for extending the capabilities of blind people. In *Extended abstracts of the 2020 chi conference on human factors in computing systems* (CHI EA '20), 1–4. Honolulu, HI, USA: Association for Computing Machinery. DOI: 10.1145/3334480.3383142. https://doi.org/10.1145/3334480.3383142.

Huang, Yueh-Min, Rustam Shadiev, Wu-Yuin Hwang & Nian-Shing Chen. 2014. Review of Speech-to-Text Recognition Technology for Enhancing Learning. en. *Educational Technology & Society* 17. 21.

Mboshi, Nsagha Sarah. 2018. TEACHING LEARNERS WITH VISUAL IMPAIRMENT IN AN INCLUSIVE EDUCATION SETTING: THE CAMEROON PERSPECTIVE. en. *International Journal of Education and Research* 6(2). 10. https://www.ijern.com/journal/2018/February-2018/11.pdf.

Morrison, Cecily, Edward Cutrell, Martin Grayson, Elisabeth RB Becker, Vasiliki Kladouchou, Linda Pring, Katherine Jones, Rita Faia Marques, Camilla Longden & Abigail Sellen. 2021. Enabling Meaningful Use of AI-Infused Educational Technologies for Children with Blindness: Learnings from the Development and Piloting of the PeopleLens Curriculum. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility*. New York, NY, USA: Association for Computing Machinery. DOI: 10.1145/3441852.3471210. https://doi.org/10.1145/3441852.3471210.

Patra, Ganesh & Subhash Chander. 2021. Education of Children with Disabilities: Exploring Possibilities with Artificial Intelligence. en. *Pedagogy of Learning* 7(3). 29–35.

Roach, John. 2018. *AI technology helps students who are deaf learn*. en. https://blogs.microsoft.com/ai/ai-powered-captioning/ (4 January, 2022).

Robles Bykbaev, Vladimir, Eduardo Pinos Velez & Paola Ingavelez Guerra. 2011. An educational approach to generate new tools for education support of children with disabilities. en. In *International Conference on e-Education, Entertainment and e-Management*, 80–83. Bali, Indonesia: IEEE. DOI: 10.1109/ICeEEM.2011.6137847. (6 January, 2022).

Sparkes, Indiana. 2019. Disability and education, UK. en. *Office for National Statistics*. 15. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/disability/bulletins/disabilityandeducationuk/2019.

Universal Declaration of Human Rights. 1948. en. https://www.un.org/sites/un2.un.org/files/udhr.pdf.

Wu, Jun. 2019. Artificial Intelligence Is Providing Special Education Alternatives. en. *Medium*. 6. https://medium.com/artificialciti/artificial-intelligence-is-providing-special-education-alternatives-389b551d16d0 (4 January, 2022).

Zeng, Fan-Gang. 2017. A New Landscape for Hearing Aids: en. *The Hearing Journal* 70(12). 6. DOI: 10.1097/01.HJ.0000527871.60334.ad. http://journals.lww.com/00025572-201712000-00002 (6 January, 2022).

Chapter 17

AI-typical learning buddy - A literature review of AI-based learning applications for people in the autism spectrum

Mara Rehmer & Katharina Trant

Autism Spectrum Disorder (ASD) is a medical condition related to brain development that includes a variety of challenges for those affected. Therefore, Artificial Intelligence (AI) methods are integrated in educational mobile applications to provide individualized support. In this literature review, we aim to provide an overview of current AI based applications for people with ASD. In the 10 papers we selected, we found that most apps are designed for children and help with verbal communication. Further emphasized are applications for emotion detection and expression. For description of their AI methods, most papers refer to earlier works resulting in minimal explanation thereof. The researchers already stated potential improvements which include studies with more participants as well as expanded international collaborations. In future works, the focus should lie upon creating a smart mentor supporting the learners and caregivers.

Keywords: Learning Apps, Autism Spectrum Disorder, AI Applications

1 Introduction

Autism Spectrum Disorder (ASD) is a medical condition related to brain development. Besides autism, the spectrum also includes other conditions like Asperger's syndrome. In the following, the whole spectrum is implied when using the term autism. The first symptoms usually arise in early childhood and include e.g. little eye contact or no response to name calling. With time other complications may emerge: problems with social interaction and communication are the most prominent ones, yet the pattern of behavior is very individual (*Autism spectrum disorder* 2022). There is no cure for this condition which is why early treatment with different kinds of therapy can significantly improve the quality of life. But therapy can be quite expensive (Tang et al. 2016).

Autism is not labeled as a learning disability, but learning can nevertheless be challenging for autistic people. Here, it can be particularly advantageous that autistic people, especially children, are found to be highly technophilic. This can be very helpful to improve learning skills in modern times (Najeeb et al. 2020). For that reason, using Artificial Intelligence (AI) based applications that target special needs and deficiencies could be ideal in assisting people diagnosed with ASD. These apps have the potential to adapt to the personal requirements, can be easily accessible if made for smartphones or computers and are less expensive than usual therapy (Tang et al. (2016); Zhang et al. (2021)).

Our interest in this topic sparked while watching the Netflix Original show Atypical, dealing with the experiences of an autistic teenager. The technical progression and especially the current pandemic situation give rise to the importance of individual and home-based learning opportunities for people with ASD who often show an affinity to work with technology (Najeeb et al. 2020). This is why our systematic literature review aims to give an insight into some of the recent AI-based applications aiding these individuals. In this review, we only include apps that are for learning and improving certain skills, but we leave out intelligent applications that help with diagnosis of ASD or structuring daily routines.

In Section 2 we describe our methodology of choosing the papers with inclusion and exclusion criteria. After that, in section 3, we summarize our findings. This section is divided into applications for autistic children or adults and applications that are targeted for parents, teachers or caregivers of people with ASD. We also name limitations and recommendations that future applications should consider. Section 4 provides our conclusion alongside a listing of the mentioned apps and a short description in Table 1.

2 Methodology

In this section we explain our search strategy and describe our inclusion and exclusion criteria along with the tools that helped us sort the information. To collect papers for our literature review we mainly focused on the Scopus database to get a comprehensive overview of relevant papers. With the keywords mentioned below we gathered 80 papers out of all results that generally seemed to fit our topic.

- "smart App for people with ASD"
- "AI application for people with ASD"
- "mobile application for people with ASD"
- "Autism spectrum disorder, AI, application"
- "ASD, Artificial Intelligence, mobile App"

After reading every abstract and the keywords of these papers, we had to exclude a large amount since many of the applications did not use AI. Furthermore, we excluded papers that were not written in the past 5 years, ones that solely focused on diagnosing ASD as well as applications targeted to caregivers to organize daily lifestyle events (e.g. a digital planner). We narrowed it down to 12 papers, but while reading them thoroughly we had to exclude two more papers. One introduced a smart toy but not an application. The other one was a review and only mentioned one relevant paper for us, which was already part of our literature. Finally, we were able to limit our collection to 10 papers in total. Rehman et al. (2021) is a review on 25 AI-based apps for autism which were selected from the app store based on high rating and user review sentiment analysis. Based on their descriptions, 14 of these apps suited our inclusion criteria. For the sake of avoiding repetitive descriptions, we decided not to include these apps in our paper since a review of them can be found in Rehman et al. (2021). For organizing and collaborative work we used Qiqqa, a free and open-source software that allowed us to work with and categorize the papers simultaneously.

3 Results

In this section we will describe applications designed specifically for people with ASD, their purposes as well as the types of AI that are used. Additionally, we shortly elaborate on applications that are targeted towards the caregivers, professionals and parents who work or live with people on the spectrum. Finally, we give a short overview on what was recommended to be changed or invented in future projects.

3.1 Applications for people with ASD

We found that the common opinion among all papers is that the applications must be highly customizable due to the variety of needs of autistic people. Consequently, there is a huge diversity in provided applications, their design and their purposes.

Every case of autism is unique, yet each individual case seems to involve some kind of emotional instability and mediocre social skills (Garcia-Garcia et al. 2019). Since one prominent autistic hardship appears to be social interaction, especially verbal communication, the majority of apps we found tries to provide some kind of communication helper. To be specific, six papers focus on improving communication skills. Zhang et al. (2021) designed an intelligent agent programmed into a serious game (i.e., game used for educational purposes) for children that is capable to communicate with the user and simultaneously measures collaboration and verbal-communication skill of the children. This collaborative puzzle is based on tangram games where two interacting players, the autistic child and the AI-based agent, have to finish the puzzle together whilst communicating with each other to exchange game information, like size or color of the puzzle pieces. Therefore, it is pivotal that the AI-agent is capable to understand unrestricted language and is able to react accordingly. Machine Learning (ML) methods help the agent to read input in real time as well as playing the game in two different modes. The first possible mode is the 'turn-taking' configuration. This enables it to monitor the child's input and waiting its turn. If the child does not make a move, the agent will prompt a question like "This is a turn-taking game. It is your turn to move a puzzle piece." (Zhang et al. 2021). In the second configuration 'move-together', the child has to verbally communicate which puzzle piece they want to move. After that the agent has to confirm the selection and they move the puzzle piece together. If a human user happens to move puzzle pieces without talking to the agent beforehand, the agent will prompt a verbal message like "Which piece should we move?"(Zhang et al. 2021). These are just a few functions of this application. The agent is also able to answer questions, provide assistance as well as reiterate the current game's rules.

Another important way to incorporate AI in educational applications for people with ASD is the use of chatbots to help communicate with autistic people as well as helping them communicate with their surroundings. Two out of our ten papers deal with AI-based chatbots. The chatbot Aliza from Najeeb et al. (2020) helps autistic children to train their speech development skills like learning the alphabet. Aliza is integrated into a smart mirror and holds many more functions than just the chatbot which are described below. The second example of a chatbot is Alex, designed by Cooper & Ireland (2018). This application helps people on the spectrum communicating with their surroundings as well as their caretakers and therapists connecting and understanding them. Alex and Aliza are applications using Augmentative and Alternative Communication (AAC). In fact, five out of our ten papers use AAC which is deemed to be very beneficial as it is based on visual symbols as well as speech-generating and can be personalized to every individual need (Cooper & Ireland 2018). According to Farzana et al. (2020), AAC "can be regarded as an approach that integrates tools and strategies (gestures, symbols, speech generating devices) to cope with daily communication challenges". Chatbots allow autistic people to train their communication in a non-judgmental environment and without the need to work towards a specific goal in a dialogue (Cooper & Ireland 2018).

Similarly, Porayska-Pomsta et al. (2018) designed an intelligent agent named Andy which, in cooperation with a person, is able to support autistic children with their communication. Andy is able to interact verbally, non-verbally or combining both and runs on a 42-inch multi touch screen which enables the children to interact with Andy in a physical and non-physical way. This application focuses on especially low functioning ASD children, to assist with absent or inappropriate responses as well as initiating a dialogue. Andy provides positive feedback, simple language and precise instructions. Porayska-Pomsta et al. (2018) found a significant increase in children's responses not only to the AI-agent but also the human social partner, indicating a great benefit of combining AI-based training with real-life training with humans, creating a good support system.

People on the spectrum may also experience shortcomings dealing with non-verbal communication or social understanding.They can, for example, have problems with identifying facial expressions or showing emotions (*Autism spectrum disorder* 2022). This is why some of the applications include facial recognition algorithms designed to detect emotions. The "Expressiveness Mirror" (Begel et al. 2020) is an app using Computer Vision (CV) to recognize the user's emotions and display them as an emoji to them. The idea is that the user, an autistic person, sees how a neurotypical person (someone who has no cognitive impairment) would conceive their facial expressions during a video call. This can potentially prevent misunderstandings, however, the reactions from the user study were predominantly negative, stating that it didn't match their own impressions and it distracted from the conversation. Additionally, there were only 6 basic emotions (happiness, sadness, fear, surprise, anger and disgust) implemented, yet the most problematic factor was detecting more complex feelings or expressions like stress, confusion or sarcasm, which are harder to implement. Begel et al. (2020) also found out from the users that they would benefit from an application that the paper named "Expressiveness Prosthetic", which reads the emotional state of the conversation partner.

Garcia-Garcia et al. (2019) developed a serious game for smartphones called "EmoTEA", which does not only use emotion recognition techniques but also Tangible User Interfaces (TUI). A TUI incorporates digital information in a physical environment that can be manipulated and explored. This can improve the motivation to interact with the system and benefit the learning process. In this case, the TUI are hard plastic cards with either pictograms or pictures of real people showing different emotions. The cards use Near Field Communication (NFC) and are incorporated in two of three possible games. In the first game, the player is shown an emotion in the app and they must match the best fitting card and hold it close to the NFC reading device. If it is correct, the next emotion is shown. The second game uses the camera of the device to interpret the users facial expression, because the task is to mimic a displayed emotion. In the third game, the user watches a video where a specific mood is displayed and afterwards uses one of the cards again to match this emotion.

The smart tutoring system Aliza, developed by Najeeb et al. (2020), is supposed to help autistic children with basic education and consists of four components. One of them is the verbal trainer as described before, another is an attentiveness tracker for emotion detection. This tracker uses real-time visual data to adapt to the users emotions, ergo if negative emotions or distraction are detected, the task is modulated or a different task is suggested to keep the child's attention. Aliza is also a writing and math mentor. These functions teach the alphabet and numbers with tasks like drawing lines and curves or counting objects. The system monitors and evaluates the performance of the children and generates a report for parents, therapists or other caretakers.

One last important area of application is the integration of real time support in the daily activities of people with ASD. Tang et al. (2016) aimed for a context-aware system that uses information like location, time, schedule, personal preference and more to provide customized as well as real-time support for activities like cooking, cleaning. They combined the users' smartphones with special sensors to build a system that is highly customizable for various specific needs. The prompts are available in many different forms (text, images, audio, video and haptics), even the level of support and the interface can be personalized. The goal of this application is to improve life quality and the independence of autistic people. Additionally, the app is able to record the information and thereby help caregivers and professionals. Finally, this app is capable to contact a caregiver in case the app does not provide enough help. In this way it also combines AI and human assistance for creating a better support system. It is mentioned that this application might also help people with ADHD, Dementia or Alzheimer's disease (Tang et al. 2016).

3.2 Applications that help caregivers or parents of people with ASD

In our research we did not explicitly look for applications that are tailored for teachers, parents or caregivers of people with ASD, but we also did not want to exclude them and now briefly give an overview of what we found.

Sevilla et al. (2018) developed a smart recommendation system called SMART-ASD to assist caregivers with selecting mobile apps for the person with ASD which are suited for their needs. The system operates by saving the users' data into an ontology. The aforementioned application of Tang et al. (2016) is able to store activity records of the user and can thereby provide therapists with details for further behavioral analysis. Additionally, this app is able to get in contact with a caregiver enabling them to give supplementary support when the app fails to help.

3.3 Types of AI

Reading the papers, we encountered two detailed descriptions of the AI technology that was used in the apps. Najeeb et al. (2020) greatly elaborated the AI system they used, describing a Convolutional Neural Network (CNN) for symbol detection and a deep learning network for reports, a speech-to-text model as well as an attention tracker. Sevilla et al. (2018) applied changes to the existing semantic web ontology "Cloud4all" and included an accurate description on what they had to adjust for their specific needs.

In contrast to these two descriptions, Begel et al. (2020) stated they developed an "AI computer vision system to detect facial expressions" but there is no further elaboration on the components of this technology. Garcia-Garcia et al. (2019) neither described their emotion detection system as they relied on an existing technology called "Affectiva", that recognizes emotions via facial expressions. Similarly, Porayska-Pomsta et al. (2018) only spoke of an intelligent character that is able to interact verbally, non-verbally through gestures or via combination of these two, without going into detail on the technical facts. These authors also referred to earlier works that describe the implementation. Likewise, Farzana et al. (2020) and Zhang et al. (2021) talked about "machine learning methods", while the latter used them in combination with Natural Language Processing (NLP) for speech-recognition, speech-generation and text-to-speech functions in their applications which were defined in prior papers. As other implementations of chatbots, Alex from Cooper & Ireland (2018) worked with the "Artificial Intelligence Mark-up Language (AIML)", which is used for "case-based reasoning and textual pattern matching". Lastly, Tang et al. (2016) talked about a "context-aware assistive system" which is capable of automated decision making on the basis of collected data.

3.4 Limitations

While the idea of using AI for applications seems promising and we saw many brilliant ideas to cope with the need for very individual solutions, there are still some obstacles that need to be taken into account. The studies we found mostly involved a scarce number of participants. Garcia-Garcia et al. (2019), for example, have only 3 children to test their app. They are aware that this is not enough and plan to conduct a study with more participants in the future. Contrary to that, Najeeb et al. (2020), collected data from 100 children, but mostly focus on the performance of their system and only touch briefly on whether Aliza actually supports the learning process of the children.

The mentioned authors experienced limitations in the data collecting process due to the global pandemic, because the participants were required to wear a mask which impedes correct classification of facial expressions. Begel et al. (2020) also mention different problems that arise with data, namely that their facial expression detection model was trained with data from neurotypical people, while it was then used on their ASD participants. But there are few other options considering that there is only little data available from non-typical populations.

Farzana et al. (2020) point out that a lot of research in this area is done mainly in wealthier countries, although ASD is about equally prevalent in all countries. Consequently, they suggest collaboration between researchers from different countries so that children in developing countries can have access to and profit from these promising inventions. It is therefore also necessary that existing applications are extended and come in various languages.

Generally, we found that the papers have good intentions, yet not all AI systems are sophisticated and accurate enough to be truly useful.

4 Conclusion

In this review, we wanted to assess the current development of applications tailored to support education for autistic people as well as their potential for improvement. We found that most of the apps are designed for children and help with verbal communication. Apart from that, non-verbal communication, emotion detection and expression play an important role in this research area.

Considering the usage of AI for these applications, we observed that only a few authors go into details about the technical background of their systems. That may be because they built their applications upon earlier designs or their focus lies on ends and not the means. Generally, they use Machine Learning models in combination with CV or NLP to suit their purposes.

To improve future development in this field, research has to incorporate more participants to both enhance the performance of their systems and to validate whether their technology actually improves the learning curve of the user. Additionally, it is crucial to increase transnational and interdisciplinary collaboration to guarantee access also in developing countries. Finally, detailed information on the technical background of the AI systems would be beneficial for future researchers.

It is important to emphasize the role of balance between AI technology and humans in order to support children's learning success as well as the necessity to modulate the learning environment individually (Porayska-Pomsta et al. 2018). We think the overall goal of such technology should be to create a digital mentor who, alongside the human caregiver and professionals, acts as a proper support system for people on the spectrum.

Table 1: Overview of mentioned apps

App	Description
Alex (Cooper & Ireland 2018)	AAC chatbot app to help communicate with pictograms
Aliza (Najeeb et al. 2020)	Gamified smart mirror consisting of a writing and math tutor, a verbal trainer and an attentiveness tracker for ASD children basic education
ECHOES (Porayska-Pomsta et al. 2018)	Touch Screen App with virtual character which encourages social interaction and communication in children
EmoTEA (Garcia-Garcia et al. 2019)	Serious game to help with identifying and expressing emotions
Expressiveness Mirror (Begel et al. 2020)	AI computer vision system to detect user's facial expressions in a video call and label them to help with communication
ICON2 (Zhang et al. 2021)	Computer based collaborative puzzle game with intelligent agent that communicates with and measures verbal-communication skills of the user

App	Description
LIVOX (Farzana et al. 2020)	AI-based android mobile app recommending pictograms based on location and time to help with communication
no name (Tang et al. 2016)	Ambient intelligence-based app to assist with daily activities
SMART-ASD (Sevilla et al. 2018)	recommendation system for parents and professionals to help select adequate technology for patients with ASD

References

Autism spectrum disorder. 2022. https : / / www . mayoclinic . org / diseases - conditions/autism-spectrum-disorder/symptoms-causes/syc-20352928 (6 January, 2022).

Begel, Andrew, John Tang, Sean Andrist, Michael Barnett, Tony Carbary, Piali Choudhury, Edward Cutrell, Alberto Fung, Sasa Junuzovic, Daniel McDuff, Kael Rowan, Shibashankar Sahoo, Jennifer Frances Waldern, Jessica Wolk, Hui Zheng & Annuska Zolyomi. 2020. Lessons learned in designing AI for autistic adults. *ASSETS 2020 - 22nd International ACM SIGACCESS Conference on Computers and Accessibility.* DOI: 10.1145/3373625.3418305.

Cooper, Anabelle & David Ireland. 2018. Designing a chat-bot for non-verbal children on the autism spectrum. 63–68. DOI: 10.3233/978-1-61499-890-7-63.

Farzana, Walia, Farhana Sarker, Quazi Delwar Hossain, Tom Chau & Khondaker A. Mamun. 2020. An evaluation of augmentative and alternative communication research for ASD children in developing countries: benefits and barriers. *Communications in Computer and Information Science* 1294. 51–62. DOI: 10.1007/978-3-030-60703-6_7.

Garcia-Garcia, Jose Maria, María Del Mar Cabãero, Victor M.R. Penichet & María D. Lozano. 2019. EmoTEA: teaching children with autism spectrum disorder to identify and express emotions. *PervasiveHealth: Pervasive Computing Technologies for Healthcare.* DOI: 10.1145/3335595.3335639.

Najeeb, R. S., J. Uthayan, R. P. Lojini, G. Vishaliney, Jesuthasan Alosius & Anjalie Gamage. 2020. Gamified smart mirror to leverage autistic education - Aliza. *ICAC 2020 - 2nd International Conference on Advancements in Computing, Proceedings.* 428–433. DOI: 10.1109/ICAC51239.2020.9357065.

Porayska-Pomsta, Kaska, Alyssa M. Alcorn, Katerina Avramides, Sandra Beale, Sara Bernardini, Mary Ellen Foster, Christopher Frauenberger, Judith Good, Karen Guldberg, Wendy Keay-Bright, Lila Kossyvaki, Oliver Lemon, Marilena Mademtzi, Rachel Menzies, Helen Pain, Gnanathusharan Rajendran, Annalu Waller, Sam Wass & Tim J. Smith. 2018. Blending human and artificial intelligence to support autistic children's social communication skills. *ACM Transactions on Computer-Human Interaction* 25. DOI: 10.1145/3271484.

Rehman, Ikram Ur, Drishty Sobnath, Moustafa M. Nasralla, Maria Winnett, Aamir Anwar, Waqar Asif & Hafiz Husnain Raza Sherazi. 2021. Features of mobile apps for people with autism in a post covid-19 scenario: current status and recommendations for apps using AI. *Diagnostics* 11. DOI: 10.3390/DIAGNOSTICS11101923.

Sevilla, Javier, J. Javier Samper, Gerardo Herrera & Marcos Fernández. 2018. SMART-ASD, model and ontology definition: a technology recommendation system for people with autism and/or intellectual disabilities. *International Journal of Metadata, Semantics and Ontologies* 13. 166–178. DOI: 10.1504/IJMSO.2018.098395.

Tang, Ziying, Jin Guo, Sheng Miao, Subrata Acharya & Jinjuan Heidi Feng. 2016. Ambient intelligence based context-aware assistive system to improve independence for people with autism spectrum disorder. *Proceedings of the Annual Hawaii International Conference on System Sciences* 2016-March. 3339–3348. DOI: 10.1109/HICSS.2016.417.

Zhang, Lian, Ashwaq Z. Amat, Huan Zhao, Amy Swanson, Amy Weitlauf, Zachary Warren & Nilanjan Sarkar. 2021. Design of an intelligent agent to measure collaboration and verbal-communication skills of children with autism spectrum disorder in collaborative puzzle games. *IEEE Transactions on Learning Technologies* 14. 338–352. DOI: 10.1109/TLT.2020.3029223.

Chapter 18

Using AI for enhancing communication skills and strategies of children with autism spectrum disorder

Lisa Artman, Ilona Martynenko & Liva Zieba

One of the main challenges that people with Autism Spectrum Disorder (ASD) face is communication. However, it is one of the main aspects of education, as one has to interact with the teacher and often also with other students. To enhance the ability of learners with ASD to interact with peers, different techniques involving Artificial Intelligence (AI) can be used. In addition, they can help teachers to choose correct communication strategies for their neurodivergent students. The aim of this paper is to provide a systematic literature review of papers on the usage of AI techniques to enhance communication skills or strategies for children with ASD. This paper concentrates on eight out of initially identified 9210 results and provides an analysis on geographical distribution of such papers and affiliations of their authors. Moreover, an in depth review of five different methods of AI that are used to work on communication skills and strategies of children with ASD is provided. Namely, those methods are: smart glasses, robots, machine learning, virtual agents and monitoring systems. In conclusion we criticize the lack of studies on the topic as well as small participant sizes used for them. We also state the potential of the research on the topic.

Keywords: Autism Spectrum Disorder, Education, Special Needs Education, Communication

1 Introduction

Autism Spectrum Disorder (ASD) is defined by CDC (2020) as a developmental disability. People with ASD usually face challenges in connection to social interaction, communication and social reciprocity as well as demonstrating restrictive and repetitive behavioral patterns (American Psychiatric Association 2013: 31).

The topic of communication skills of children with ASD is not only important regarding the children's social development, like building and maintaining a social circle, but it also surrounds the scope of education, as learning challenges are one of ASD's symptoms.(CDC 2021) In that regard, it has been found that people with ASD have difficulties not only in school but also in their following professional life. At least half of young adults with ASD are not employed and have many difficulties starting with their higher education and career for the first two years after high school (Shattuck et al. 2012).

In order to make higher education homogeneous to all children, - neurotypical (meaning all people who do not show atypical neurological patterns or behavior) and not - K-12 education has to be improved for children with special needs. We believe that one of the possibilities to provide quality education that caters to the needs of children with ASD is by using different methods of artificial intelligence (AI). This paper presents a systematic review of literature on the topic of using AI as a way to enhance communication skills and strategies of children with ASD. This review focuses on papers that directly deal with the effect of different AI technologies on communication aspects in education of children with ASD. Specifically, this paper aims to answer the following questions by means of a systematic review of literature:

- How does implementation of AI techniques influence the quality of communication strategies?
- Can AI help to enhance the communication abilities of children with ASD?
- Can AI help the teachers to create and implement better communication strategies for children with ASD?

In contrast, this review does not include papers that deal with following topics: the usage of AI in teaching specific individual subjects; other learning/teaching strategies; and enhancing communication strategies for neurotypical children.

2 Methodology

This section presents methods and search strategies that were used to choose papers for this systematic review. Moreover, we defined six exclusion and inclusion criteria to choose the papers, which we introduce in Table 1.

We have mainly used the Google Scholar database to get most of the papers for this review. The search was conducted in December 2021 and was performed using the advanced search function of the Google Scholar Database. There we specified in the field “Find article” in subsection “with all words” following keywords: “Artificial Intelligence, AI, Autism, ASD, Education, Learning, Communication” and in the subsection “with at least one of the words”: “Study, ”Communication skills”, ”Social Communication”. This search yielded 9210 results, which were sorted by relevance by Google Scholar. We have chosen the first 200 of them for further consideration. We then removed 10 search results that were published before 2011. Subsequently we removed 13 duplicates. Further 53 papers were removed due to the lack of connection to ASD/Education/Artificial Intelligence and overall discrepancy from inclusion criteria (see Table 1). Moreover, 28 reviews and other non-primary sources were removed. We were therefore left with 104 papers. We read all of the titles. From the ones that seemed most fitting for our topic, we also read the abstract. We then chose five papers for a detailed analysis in our review, since we found them to be the most interesting and relevant from the information in their abstracts.

To make the selection of AI techniques more diverse, we have also utilized the search function on websites of Springer and IEEE Explore. There we have chosen three further papers for this review using the same inclusion criteria from before (see Table 1).

Table 1: Inclusion and exclusion criteria

Inclusion Criteria	Exclusion Criteria
Published between 2011-2021	Published before 2011
Published in English language	Published in other languages
Primary research	Not primary research
Research deals with Communication	No connection to Communication
Methods of AI are considered	Methods of AI are not considered
Children with ASD are considered in research	No connection to children with ASD

Our search strategy, however, includes certain limitations. We have only used one database, namely Google Scholar to perform the searches (other databases were just briefly used). Due to time constraints and the scope of the seminar within which this paper is written, we considered only the first 200 results from the database, so we may have missed some potentially fitting papers. Due to the aforementioned reasons we have only analyzed eight papers for this review, so they might not represent the full picture of the field. Moreover, the structure of keywords might have been not the most fitting one, therefore not allowing us to find other thematically passing papers.

3 Results

This section presents the results of our findings. It is further divided into subsections. The first two subsections concentrate on information about geographical distribution and the authors of the papers while the remaining five introduce papers from different sub-fields within artificial intelligence that have been used to enhance communication skills or strategies of children with ASD.

3.1 Countries

To get an idea about how the papers are distributed around the world, we looked at the countries where research was performed. We discovered that our sample of eight papers is rather diverse geographically and includes articles from six countries. The USA and the UK contributed the most papers (25% each). The full distribution can be seen in Table 2.

Table 2: Geographical distribution of papers

Country	Number of papers	Percentage, %
The USA	2	25
The UK	2	25
Bangladesh	1	12.5
Malaysia	1	12.5
Portugal	1	12.5
Australia	1	12.5

3.2 Author affiliations

Then we looked at what backgrounds and affiliations the authors that published the selected papers have. Only the first author of each paper was taken into consideration. Authors with a background in Science (including psychology), Technology, Engineering and Mathematics (STEM) contributed to six (75%) of the papers, three of which were written by people with an affiliation to Engineering. The full distribution can be seen in Table 3.

Table 3: Author affiliations

Affiliation	Number of papers	Percentage, %
STEM	6	75.0
Education	1	12.5
Medicine	1	12.5

3.3 Smart Glasses

Smart glasses are one of the tools to make use of when trying to improve communication skills. Smart glasses are a wearable technology, which consists of a display in front of the wearer's eye to insert specific data in their visual field. They can also be connected to an audio device to transmit auditory information.

Sahin et al. (2018) published the pilot study "Augmented Reality Intervention for Social Communication in Autism in a School Classroom [...]". This three week longitudinal study included one participant who was 11 years old and already had previous interventions. Smart glasses were employed by using the Empowered Brain, a tool that consists of smart glasses and selected software modules. These modules coach the wearer in facial emotion recognition, that helps with crossing over into different environments, and improves their overall social attention. The tool utilizes a special model for social communication intervention. First, the participant wore the glasses for one week without any interventions taking place, so that his ordinary behavior could be captured. Following this, in the next two weeks interventions were arranged twice a day (16 in total, 4 school days per week), each of which lasted 10 minutes. During these sessions, the participant was asked about his current day in school. In eight of the ten conversation minutes, the participant wore the Empowered Brain utensil. The primary observers concluded that the study's results show enhanced social motivation, cognition and communication. Additionally, the Social Responsiveness Scale (SRS - Second Edition), which is a scale that measures the social behavior of people with ASD, improved. However, a more distant professor did not notice any substantial changes (Sahin et al. 2018).

Another use of smart glasses was introduced in an article by Daniels et al. (2018), which describes a research documenting the experience and benefits for 14 participants (aged 9.57 on average) with ASD who were using Superpower Glass, a machine-learning-assisted software system developed by the authors of the paper. Daniels et al. (2018) made the study with the intention to evaluate the potential of Superpower Glass as a wearable therapy tool that can help to enhance social skills, eye contact and emotion recognition for children with ASD.

Superpower Glass System consists of Google Glass that can be worn by the participant and connected to an app on the Android smartphone. Google Glass is able to record the social interaction of the child and send the facial tracked data to the app. The app uses that data to compute the emotion of a person interacting with a child and returns it on the screen in the form of a social cue (represented by an emoticon and color as well as recognized audio of expression). The app also includes several game modes that allow the child to look back at the recordings of previous interactions and use them to train on recognizing correct emotions (Daniels et al. 2018). The system was used at home during three 20-minute sessions per week by the participants for 72 days on average. According to Daniels et al. (2018) after this experiment was over, participants' families reported positive results such as an increase of eye contact of the child in 85.7% and increased emotion recognition in 78.6% of the cases. Moreover, the score of Social Responsiveness Scale (SRS-2, higher SRS-2 score represents higher severity of ASD) of participating children was measured before the experiment and after. While the mean of total SRS-2 score was 80.07 points at the intake meeting, it decreased by 7.38 points on average at the conclusion appointment. Significant decrease in total SRS-2 score was also shown by ANOVA analysis. In addition, for six participants the severity class of ASD changed from more to less severe one.

3.4 Robots

Another niche of artificial intelligence is human-robot interaction. Robot-based interventions are an approach to improving social and communication skills of children with ASD.

The first study we found that made use of humanized robots was conducted by Shamsuddin et al. (2015). This pilot study presents how robot-based interactions can develop better communication skills using scenarios from the Applied Behavior Analysis (ABA). The humanoid robot had different modules of interaction – in the beginning simple modules were used, and relative to the children's mastering of them more complex modules were utilized. In the preliminary study 10 out of 12 children showed reduced autistic traits when interacting with the robot (comparing the behavior to the Autism Diagnostic Observation Schedule) and were more open to communicating with the robot. The main goal of the study was to find out whether communicating with robots is easier for children with ASD than communicating with people. And indeed, one finding of the study was that autistic children are more prone to communicate with robots as they do not judge the children's behavior and responses, and seem simpler appearance-wise. Also, the conclusion of this study improved the procedure of using a robot when enhancing children's communication. These enhancements can be used in special education schools where a similar model has been developed that also uses the ABA approach. The study concludes that robots are expected to increase the autistic children's engagement in learning and that its actual effectiveness will be evaluated with the results of trials in schools.

Another study we found wanted to promote social interaction and communication. Costa et al. (2011) released a paper, where the only participant was an 11-year-old child. It is important to include the fact that this child was nonverbal but could produce vocalizations and that they had difficulties not only in social interactions but also with keeping attention. The activity consisted of a researcher interacting with a robot by rolling a ball back and forth and the hope was that the child would verbally request to play along. Over two months the child was first familiarized with the researcher and with the exercise including the robot, before the test stages started. In the end, the child managed to perform the task much better than in the previous pre-test stage. Also, the child played for a long time with the robot as well as known and unknown partners, which is a success in itself. (Costa et al. 2011)

3.5 Machine Learning

Since Machine Learning algorithms are able to figure out patterns from data (Russel & Norvig 2010: 2), and therefore allow to map the behavioural patterns of children with ASD for future analysis and as additional information to teachers, they also present a valuable tool for special needs education.

In the research by Lampos et al. (2021), over 5000 observations of interactions between teachers and seven children with ASD were done. Lampos et al. (2021) concentrated on observing student's emotional state, teacher's communication strategies or pairs of such, and student's degree of response to them along with collecting information about students' attributes (age, sex, P-level[1], SCERTS[2]). Then, a machine learning algorithm was used to predict students' response to a particular teacher communication strategy either with or without taking into account the students' attributes. Using this kind of classification, Lampos et al. (2021) found out that visual communication strategies tend to generate the best

[1]P-level (P-scales) includes an assessment of language, mathematical and science competences of learners with special educational needs in England (Department for Education (DfE) 2017)

[2]SCERTS is a framework that is focused on helping individuals with ASD to gain competencies related to Social Communication, Emotional Regulation and Transactional Support (SCERTS 2022). Here Lampos et al. (2021) used SCERTS to classify language competence of participants into three categories: social partner, language partner and conversational partner

response from children with ASD. It was also noted by the authors of the paper that the accuracy of the classification system when ignoring the student attributes was 0.664, but it increased by about 4.37% when taking them into account. Further increase of accuracy was noticed when past information was incorporated into the system (0.711) (Lampos et al. 2021). Therefore, authors state that the classification system can be used to aid teachers in choosing the best communication method for their students with ASD. Moreover, it allows the teacher to cater the communication strategies to each individual child (Lampos et al. 2021).

3.6 Virtual Agents

One possible way in which Artificial intelligence can help autistic people with their communication skills is by using virtual human tutors.

One of the studies concerned with that topic is called "Personalisation and automation in a virtual conversation skills tutor for children with autism" and was conducted by Milne et al. (2018). The experimental design included a treatment group and a control group, each with 16 participants. The study investigated whether the social skills of the children with high functioning autism would improve after they have trained with the virtual tutor over a time period of three months. They used the Thinking Head Whiteboard program which included a teacher character, a peer with strong social skills and a peer with developing social skills. The control group used the same software, however they did not receive social learning content. The majority of the caregivers and children reported that they liked working with the software (Milne et al. 2018). Two different tests were conducted after the training period for every participant. Parental reports and observations were used to measure generalization of skills to real-world context and a questionnaire to measure the theoretical knowledge of participants. The treatment group also showed a statistically significant improvement in their social ability. Whereas the control group did not show any significant change (Milne et al. 2018). However, it seems like further studies need to be done to validate the results. This is one of the biggest and most relevant studies with regard to the topic. In our opinion, most of the other studies are either too small or do not follow a research design with a control group and therefore, it is difficult to get statistically significant results.

Porayska-Pomsta et al. (2018) published a study that examined the efficacy of a learning environment that utilises an artificial agent used in the company of a human interactor, to help children with ASC (Autism Spectrum Conditions)[3]

[3]Terms ASD and ASC are used in our systematic review interchangeably

develop their communication skills. The learning environment used in this study is called ECHOES, which is a single-user technology that uses an artificial agent called Andy to interact with the children. ECHOES gives users an opportunity to explore different situations and to rehearse them multiple times. Social communication, in particular joint attention and symbolic use, was the prime focus in learning activities for interactions between children and the ECHOES system (Porayska-Pomsta et al. 2018). Main questions of the research were whether ECHOES enhances the response and/or initiations of social interactions for children with ASC as well as whether this increase may be transferred to other contexts. 15 children with ASD participated in the full study process that consisted of firstly familiarisation and observation from the researchers and then pre-test table-top activities, ECHOES intervention and post-test table-top activities with the participants. Most of the activities during the learning sessions with ECHOES were performed by Andy and the child cooperatively. Out of the two activities where the child was the only participant, at least one was used at the beginning of each learning session during the experiment. As a result of this research, the proportion of responses of children with ASC during the interactions in the ECHOES environment increased compared to the pre-test table-top activity (0.814 as opposed to 0.656). However, this increase has not transferred outside of the ECHOES environment, as the proportion of responses equaled 0.711 in post-test table-top activities. Also, there was no significant increase in initiations of participants to the social partner. However, qualitative data collected and observations still suggest the possibility of transfer from ECHOES to the classroom environment, as some anecdotal evidence showed the improvement of communication abilities of individual participants (Porayska-Pomsta et al. 2018).

3.7 Monitoring Systems

Lastly, we want to take a look at a slightly different approach. As we mentioned earlier, social interaction and communication is challenging for people with ASD (American Psychiatric Association 2013: 31). Therefore monitoring systems can help individuals with ASD to communicate their emotional state and get the support they need.

Al Banna et al. (2020) published a study, that took a look at how monitoring systems can improve the life of autistic people. Children with ASD were given a smart wrist band with a variety of different sensors to monitor their current state of being. Furthermore, a camera recorded their facial expressions, as long as they were sitting in front of their computer screen. The data was then analyzed by an artificial intelligence and caregivers were alerted if any abnormalities occurred.

4 Conclusion

Overall, we have analyzed eight papers on the topic of artificial intelligence in connection to communication skills and/or strategies for children with ASD. Within our small sample set, we discovered that some of the patterns could still be recognized, such as both the USA and the UK producing the lead amount of papers. Moreover, the papers were mostly written by authors with affiliation to STEM.

During our search, we have noticed the lack of papers and research dealing precisely with communication skills and/or strategies for children with ASD in context of using AI. However, out of research which focused on using technologies for improving communication strategies for learners with ASD, quite a lot were working with smart glasses or robots. We think that one of the reasons for such popularity of smart glasses might be due to the availability of technology.

Another pattern that we noticed is the rather small size of participants used for researches, ranging from one up to about 15 participants. The frequently found single-subject researches leave a less prominent generalizability, which is why it would be an enhancement to do research with bigger groups of participants.

This paper concentrates on eight papers that deal with five AI tools and methods (smart glasses, robots, machine learning, virtual agents and monitoring systems). We think, that the findings of authors of the papers could improve communication skills of children with ASD and be of great help for teachers. In our opinion, there is a hopeful future for research in this direction. Using modern techniques helps children with ASD to enhance their communication skills. In addition, such methods of AI, in particular machine learning, are a great tool for teachers who work with learners with ASD, since it allows them to ensure best communication methods with individual students. This in turn improves the quality of education for autistic children.

To conclude, even though there is the lack of research in the field, the number of papers in the last few years has significantly increased, which leaves us hopeful for the future of research. The work that has been conducted already had significant benefits for children with ASD, especially in education, so enhancing the use of AI and further research is crucial.

References

Al Banna, Md Hasan, Tapotosh Ghosh, Kazi Abu Taher, M Shamim Kaiser & Mufti Mahmud. 2020. A monitoring system for patients of autism spectrum disorder using artificial intelligence. In Mufti Mahmud, Stefano Vassanelli, M. Shamim Kaiser & Ning Zhong (eds.), *Brain Informatics*, 251–262. Springer International Publishing. DOI: 10.1007/978-3-030-59277-6_23.

American Psychiatric Association. 2013. *Diagnostic and statistical manual of mental disorders*. 5th edn. Arlington, VA: American Psychiatric Publishing.

CDC. 2020. *What is autism spectrum disorder?* https://www.cdc.gov/ncbddd/autism/facts.html.

CDC. 2021. *Signs and symptoms of autism spectrum disorders*. https://www.cdc.gov/ncbddd/autism/signs.html.

Costa, Sandra, Filomena Soares, Cristina Santos, Manuel J Ferreira, Fátima Moreira, Ana Paula Pereira & Fernanda Cunha. 2011. An approach to promote social and communication behaviors in children with autism spectrum disorders: Robot based intervention. In *2011 RO-MAN*, 101–106. IEEE. DOI: 10.1109/ROMAN.2011.6005244.

Daniels, Jena, Jessey N Schwartz, Catalin Voss, Nick Haber, Azar Fazel, Aaron Kline, Peter Washington, Carl Feinstein, Terry Winograd & Dennis P Wall. 2018. Exploratory study examining the at-home feasibility of a wearable tool for social-affective learning in children with autism. *NPJ Digital Medicine* 1(1). 32. DOI: 10.1038/s41746-018-0035-3.

Department for Education (DfE). 2017. *Performance-P Scale-attainment targets for pupils with special educational needs*.

Lampos, Vasileios, Joseph Mintz & Xiao Qu. 2021. An artificial intelligence approach for selecting effective teacher communication strategies in autism education. *NPJ Science of Learning* 6(1). 25. DOI: 10.1038/s41539-021-00102-x.

Milne, Marissa, Parimala Raghavendra, Richard Leibbrandt & David Martin Ward Powers. 2018. Personalisation and automation in a virtual conversation skills tutor for children with autism. *Journal on Multimodal User Interfaces* 12(3). 257–269. DOI: 10.1007/s12193-018-0272-4.

Porayska-Pomsta, Kaśka, Alyssa M Alcorn, Katerina Avramides, Sandra Beale, Sara Bernardini, Mary Ellen Foster, Christopher Frauenberger, Judith Good, Karen Guldberg, Wendy Keay-Bright, Lila Kossyvaki, Oliver Lemon, Marilena Mademtzi, Rachel Menzies, Helen Pain, Gnanathusharan Rajendran, Annalu Waller, Sam Wass & Tim J. Smith. 2018. Blending human and artificial intelligence to support autistic children's social communication skills. *ACM Transactions on Computer-Human Interaction (TOCHI)* 25(6). 35. DOI: 10.1145/3271484.

Russel, Stuart & Peter Norvig. 2010. *Artificial Intelligence: A Modern Approach*. 3rd edn. New Jersey: Pearson Education, Inc.

Sahin, Ned T., Rafiq Abdus-Sabur, Neha U Keshav, Runpeng Liu, Joseph P Salisbury & Arshya Vahabzadeh. 2018. Case study of a digital augmented reality intervention for autism in school classrooms: associated with improved social communication, cognition, and motivation via educator and parent assessment. *Frontiers in Education* 3. 57. DOI: 10.3389/feduc.2018.00057.

SCERTS. 2022. *Learn more about The SCERTS® Model*. https://scerts.com/.

Shamsuddin, Syamimi, Hanafiah Yussof, Fazah Akhtar Hanapiah, Salina Mohamed, Nur Farah Farhana Jamil & Farhana Wan Yunus. 2015. Robot-assisted learning for communication-care in autism intervention. In *2015 IEEE International Conference on Rehabilitation Robotics (ICORR)*, 822–827. IEEE. DOI: 10.1109/ICORR.2015.7281304.

Shattuck, Paul T, Sarah Carter Narendorf, Benjamin Cooper, Paul R Sterzing, Mary Wagner & Julie Lounds Taylor. 2012. Postsecondary education and employment among youth with an autism spectrum disorder. *Pediatrics* 129(6). 1042–1049. DOI: 10.1542/peds.2011-2864.

Chapter 19

AI in education for children with intellectual disabilities

Cosima Oprotkowitz & Isabel Grauwelman

This literature review looks into research on AI in education for children with intellectual disabilities. With the focus lying on the use of AI, including the designated users and the related difficulties, we also look into shortcomings the presented papers mention as well as general ones we find relevant. The descriptive findings show that most reviewed papers are published in journals on education and technology, while none are published in AI-specific journals. The further results show that a majority of the papers propose and/or test learning systems and use machine learning as the applied method of AI. Adaptive learning systems used by the students seem to be one promising way to incorporate AI in this context. However, the usage of AI is not clearly communicated in the papers and the proposed systems have only been tested short-term. We conclude that more research is needed to investigate how AI-systems can benefit students with IDs in their education.

Keywords: intellectual disabilities, education technology, special education

1 Introduction

Intellectual disabilities (IDs) are neurodevelopmental disorders beginning in childhood, characterized by deficits in intellectual domains, for example reasoning and deficits in the ability to adapt to developmental and social standards. Commonly known examples of IDs are Down's syndrome or the Fragile X syndrome. Often, IDs co-occur with or include other disorders, for example, learning disabilities (LDs) or communication disorders. This usually leads to a need for closer supervision and help, thereby affecting the individual's independence and in turn likely their quality of life. (Committee to Evaluate the Supplemental Security Income Disability Program for Children with Mental Disorders et al. 2015) The percentage of people with an ID in Europe is estimated to be less than one percent of the whole population, but statistics tend to differ depending on the methodology and the definitions of ID (European Intellectual Disability Research Network 2003).

Besides some possible pharmacological interventions, the treatment of IDs usually aims at improving the individual's life skills by early behavioral and cognitive therapies as well as special education, depending on the individual and their abilities (Committee to Evaluate the Supplemental Security Income Disability Program for Children with Mental Disorders et al. 2015). Further, students with LDs, which are a common part of IDs, are found to have a positive attitude and motivation towards the use of modern educational technology compared with more traditional methods of teaching (Lepičnik Vodopivec & Bagon 2016), which is a good premise for the use of technology that involves artificial intelligence (AI) in this field.

In this review, we focus on journal-published research papers dealing with the use of AI in school or related education for children with IDs. Papers solely focusing on physical disabilities or other mental disorders are not included, but can appear in some papers as those topics tend to overlap (see chapter IV for more on learners with physical impairments). The main questions we aim to answer in this review are the following:

1. What is AI used for in the context of education for children with ID?
2. Which aspects of IDs do the studies aim at?
3. Who are the designated users of the technology - students or teachers?
4. What are shortcomings of the displayed usage of AI?

2 Methodology

To gather papers for this review, we primarily used the Google Scholar search engine with its automatic sorting for relevance and filtered for results from 2010 or later to include rather contemporary research. We limited the search to the first five result pages for each of the search terms, which can be seen in Table 1. We used quotation marks to ensure that the respective terms are included in the results.

Table 1: search terms

students with intellectual disabilities ”artificial intelligence”
“intellectual disability” in education ”artificial intelligence”
”artificial intelligence” education ”intellectual disabilities”
“artificial intelligence” in special education
“learning system” learners with intellectual disabilities
”intellectual disabilities” school education ”artificial intelligence”
“intellectual disabilities” “artificial intelligence”
”intellectual disabilities” school ”artificial intelligence”
learners with intellectual disabilities ”ai”
learners with intellectual disabilities ”artificial intelligence”

From these results, we excluded papers that did not fit the following criteria: journal-published, open or institutional access for Osnabrück University, dealing with AI, children with IDs and school or related education, and written in English. This especially excluded reviews, overviews and conference papers, as well as philosophical papers and those on diagnostics of IDs, because those were not in our focus or interest. Additionally, we looked through the references of Kharbat et al. (2020)’s literature review - as it appeared as the most popular result for almost every search term - and of Patra & Chander (2021)’s literature review - as it is the most recent one - to find more relevant papers. Due to our criteria, this only led to one more result.

This whole process resulted in ten papers which were equally distributed between the authors to work on. The papers were reviewed according to previously determined aspects that are introduced below.

3 Results

In this section, we present the results of our literature review with respect to their journal types, author affiliations, geographical distribution and study types as well as the targeted aspects of IDs, the types and uses of AI, the studies’ results and their shortcomings.

3.1 Journal types

Of the ten papers, a majority of six is published in journals on education and technology. Two of these are published in the *International Journal of Information and*

Education Technology, another two in *Educational Technology & Society* and the others are published in either the *British Journal of Educational Technology* or the *Education and Information Technologies*. Another three papers are published in technology-focused journals, that is in *SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology*, *Journal of Information and Organization* or *International Journal of Social Robotics*. The remaining paper is published in an education-focused journal, namely *Educational Sciences: Theory & Practice*. Notably, none of the papers are published in an AI-focused journal.

3.2 Author affiliations

Half of the papers have been written in collaboration by differently affiliated authors, meaning that they are interdisciplinary in nature. Four have been conducted by Computer Science or Computer Engineering department members, and only one paper has been written by authors associated with Human Resource Management and Inclusion Work.

3.3 Geographical distribution of the universities

For a geographical overview, all authors' universities at the time of publication are taken into account. These are quite diverse: five out of the ten studies have been conducted in Asia (Pakistan, Qatar, Turkey and two in India), three in European collaborations (Italy and the UK; Slovenia and Germany; the UK, Spain, Italy and the Netherlands), another in a collaboration of the UK and Brazil and the remaining one in Morocco, Africa.

3.4 Study types

Regarding the methodology, four papers propose a new learning system, approach or interface and additionally evaluate and/or test it in a subsequent study with children with IDs. Similarly, two papers contain a proposal without further testing. In two other papers, the authors report qualitative studies by interviewing teachers, caregivers, experts, parents and/or the affected children themselves, while the remaining studies are experimental in nature.

3.5 Targeted aspects of IDs

Those papers mentioning their targeted aspect of IDs aim at alleviating basic language learning and mathematical concepts (Bhagchandani et al. 2020), communication and concentration (Cigale et al. 2018, Faria et al. 2020), dysgraphia, dyslexia and dyscalculia (Tariq & Latif 2016, Polat et al. 2012) or emotional or mental stress and social pressure in classrooms and related learning scenarios (Standen et al. 2020, Ouherrou et al. 2018, Conti et al. 2017). Four out of the ten papers do not specify what aspect they aim to target.

3.6 Results of the studies

Overall, the experimental studies show quite positive results concerning the engagement of the students with the proposed and tested learning systems (e.g. Tariq & Latif 2016, Saad et al. 2015). At the same time, some papers point to the complexity involved in developing such learning systems as a disadvantage of AI use in this context, because it takes a lot of time and systems would need to be adapted to different schools' curricula to allow for an in-class application (e.g. Saad et al. 2015, Faria et al. 2020). But as Polat et al. (2012) suggest, adaptive learning systems could help children with IDs or LDs to get the support they need, even when teachers and parents lack the necessary knowledge on how to educate children with IDs. Likewise, Garg & Sharma (2020) conclude that AI in general can be helpful by making inclusive school situations easier for children with special needs.

Under the premise that these students suffer from emotional stress while learning in typical classroom settings, Ouherrou et al. (2018) investigate whether Virtual Learning Environments (VLE), in this case in virtual reality, could alleviate this stress and find that children with LDs likely experience less negative emotions in VLEs. However, this is greatly dependent on the child's personality: Children with LDs who were generally curious and had a strong personality showed high motivation and engagement to complete educational games, while those that were rather ashamed and anxious got tired more easily over time and showed less desire to complete the game. Regardless of this, all children liked the animal companion, that helped them navigate through the VLE, and the multimedia approach. For a review on VR applications in special needs education, see chapter III.

Similarly, Standen et al. (2020) assess the children's affective state while learning in order to adapt the learning content accordingly. They build two tutorial systems: One solely focusing on the child's achievements for adapting the content and one that additionally considers their affective state. Comparing the learning progress of children using these systems, they conclude that additionally considering the affective states (as identified by a CNN) does not increase the systems' effectiveness significantly, but increases the engagement with the content and lowers boredom, both of which are seen as important factors for the learning progress.

Faria et al. (2020) also base their approach on the classification of mental states by having the ML-based system trained on EEG data. As their goal has been to evaluate whether children liked the system, they have not tracked any learning progress, but conducted a questionnaire on usability and acceptance only. On this, they have received consistent positive feedback.

Saad et al. (2015) and Tariq & Latif (2016) also test their proposed learning systems and have received positive feedback from both children and teachers. In Tariq & Latif (2016), the acceptance questionnaire shows that the children have got used to the application quickly and show an ”increased preference in learning” for their application in comparison to learning on ”paper”, as well as showing lower frustration levels and an increased sense of self-confidence.

Another way to incorporate AI-based systems into special education is to use robots like Conti et al. (2017) described. They investigate the acceptance of ”socially assistive robots” that practitioners can use for both educational and care assistance, and find that experienced practitioners are not very convinced to use them, while future practitioners show high interest to do so and have enjoyed interacting with the robots.

Bhagchandani et al. (2020) and Cigale et al. (2018) do not report any results, because they have not user-tested their proposed new learning systems yet.

The general agreement among the papers seems to be that gaming may be a good idea for alternative educational approaches, as all of the proposed learning systems used playful content to encourage learning (see also chapter I for another gamification approach). Additionally, virtual or on-screen learning environments seem to be able to reduce ”social frustration” in tasks as well as allowing for free exploration of actions without the perceived pressure that comes with a controlling supervisor or caretaker.

3.7 Types and applications of AI

In a substantial majority of seven papers, machine learning (ML) is used as the type of applied AI and as such, AI is only one of multiple parts in the developed algorithms and systems. Sometimes, the notion of ”machine learning” is specified, for example Ouherrou et al. (2018) use a Convolutional Neural Network (CNN) and Standen et al. (2020) use CNNs and Support Vector Machines for their adaptive learning system.

Similarly, some papers mention whether they use supervised, unsupervised or reinforcement learning (e.g. Cigale et al. 2018, Tariq & Latif 2016). Nevertheless, Tariq & Latif (2016) are ambiguous on whether reinforcement learning has been used to optimize the non-AI learning algorithm or to improve the learning content based on the user's performance. Other times, ML is mentioned but not specified further (e.g. Faria et al. 2020, Bhagchandani et al. 2020). One paper deals with AI in robots (Conti et al. 2017), while another paper simply refers to "intelligent teaching systems" (Polat et al. 2012). Another paper, aiming to analyze "AI's impact on education for students with special needs", only lists examples of AI, like robots or speech recognition, as technologies that could help children with IDs (Garg & Sharma 2020).

Reading the papers while focusing on AI, it is noticeable that only a few papers are really clear in their explanations concerning the used AI. Meaning that, overall, we find a range from good and specific explanations, like in Standen et al. (2020), Ouherrou et al. (2018), to very vague or nearly non-existent explanations, like in Saad et al. (2015), Faria et al. (2020).

The mentioned types of AI are mostly built in the respective systems as part of their algorithms or used as a helping tool to measure, for example, emotional stress (Ouherrou et al. 2018). Thus, these types of AI are always utilized indirectly when a system or method is used by a student or teacher. Out of the two papers which focus solely on proposing a new system, in one it is specified that the system is intended to be used by caregivers, while the other's authors one only state that their system is meant for "providing early intervention services" (Bhagchandani et al. 2020) for children.

In some studies, the AI is utilized for affective or mental state classification or face recognition (e.g. Ouherrou et al. 2018, Standen et al. 2020), while one interview-based study finds that systems adapting the learning content to the individual user are promising for special needs education (Polat et al. 2012), which is realized by ML in the studies proposing a system. Consequently, the overall aim of the usage of AI (in these systems) is to facilitate better and easier learning experiences by tailoring them to the user's needs.

3.8 Shortcomings

Standen et al. (2020) point out that even though an AI-based system gives promising results, "it has to be emphasized that these variables represent how the system was interpreting the affective state of the learner", thereby reminding the reader of the possible fallibility of AI-systems. They also hint at the short-term character of the study and the results respectively, as short-term testing may not reliably show the actual benefits or effects of a learning system in the long term.

Additionally, Polat et al. (2012) and Conti et al. (2017) mention that teachers, parents, practitioners and other caregivers of children with IDs need to be informed and taught about the functioning of modern learning systems in order for these technologies to be used optimally and appropriately.

One aspect, only mentioned once by Cigale et al. (2018), is the risk of a possible misuse of their system. It differs from the others in terms of use, as it gives out action proposals for an individual's caretaker based on the individual's mood as assessed by their ML-based system from data on their body language, physiological parameters and speech, whereas the other systems are designed to be used by the children or students (for teacher assistants, see chapter I). Still, the other nine papers neither address any possible misuse of the described systems or technologies, nor cover other ethical considerations when it comes to data collection or possible surveillance.

While all proposals of new learning systems either mention the individual target user or even test it with participants in an individual setting, none of them outline how the systems are thought to be utilized in a broader setting. This would include information on whether a system is designed for classroom-wide use or for individual students, for in-class or at home use, for special or standard education settings, for supervision by a teacher, parent or expert, and whether social interactions with other students affect the learning progress.

4 Conclusion

To conclude, we summarize our findings by answering our questions from the introduction.

(1) What is AI used for in the context of education for children with IDs?

If AI is mentioned as part of a study, it was mostly not explained further. It became clear that AI is mostly used as part of a proposed new learning system for children with IDs or as a tool to measure emotional states in order to get an idea on how to improve those systems.

(2) Which aspects of IDs do the studies aim at?

Overall, the few papers we found mention a diverse range of applications and targeted difficulties, but the majority does not specify what aspect of learning difficulties children with IDs experience they aim to address. Remarkably, oftentimes different disabilities or disorders are generalized, meaning that sometimes autism and ADHD are also subjects of the studies, as some difficulties hold for them as they do for IDs apparently. Studies that do focus on a single aspect target the emotional stress in classroom situations (Ouherrou et al. 2018), handwriting difficulties (Tariq & Latif 2016), concentration difficulties (Faria et al. 2020) and communication difficulties with intellectually disabled children (Cigale et al. 2018).

(3) Who are the designated users of the technology - students or teachers?

The studies that utilize AI as part of a technology that is designed for a user most often proposed systems that should be used by students as a learning system. As exceptions, Cigale et al. (2018) propose a system that is supposed to make the communication with the children easier for caregivers as users, and Conti et al. (2017) assess the acceptance of an AI-based socially assistive robot that could both be used by children alone or as a teaching assistance for practitioners.

(4) What are shortcomings of the displayed usage of AI?

No study places its focus on AI, meaning that evaluating the usage of AI in their context is almost never part of it. Mentioned shortcomings are that even an AI-system that works as intended may fail to capture the ”human reality” (Standen et al. 2020) and that it is important for designated users and assistants like teachers or parents to be properly trained to efficiently use modern technology in order for it to be applied on a larger scale (Polat et al. 2012).

5 Discussion

With our ten papers, we obviously look at a small number of papers with quite narrow criteria. One additional source that could bring more insights are conference papers, which we have not included to have a more homogeneous group of papers. There are likely also more papers that would theoretically fit our topic, but just miss one of our criteria - papers that deal with education for children with IDs without any involvement of AI can for example hint at possible use cases of AI in this context.

While conducting this review, it has quickly become apparent that there are seemingly few research papers on the specific topic, although the presented papers are not pessimistic towards the application of AI in the (classroom) education for children with intellectual disabilities. Most even point out that technology-based education methods work well for children with IDs regarding their engagement with the systems. One reason for that is that virtual or on-screen learning environments can reduce frustration in the students concerning social aspects like negative comments from other students or pressure to perform well in front of the class and teacher.

Drawing on this, further research is needed to work on learning systems for intellectually or otherwise disabled children, to lift more social and emotional pressures off of them while helping them learn. Such systems, before being applied on a larger scale, would need long-term testing in school and home situations in addition to first laboratory experiments to better show their potential and additional possible shortcomings that do not appear in a controlled environment.

To enable the use of potentially very helpful systems, the educators and caretakers, who are involved in the application and in assisting the children, would need to be educated on the systems and their functioning. This, in our opinion, would also include high-level education on the inner workings of the respective applications, e.g. explanations on the AI parts and how they roughly work and what kinds of data are gathered and used for the systems to work, to allow for an informed use. Related to that, something we would have wished for from a reader perspective is more clarity on the use and notion of AI in the papers, especially considering that potential readers are educators with no prior knowledge on technology or AI.

Further, as Polat et al. (2012) mention, such learning systems could be useful to support children with their needs in the case that teachers or parents are not equipped enough to support them properly. Still, we think that technology, even if it is functional and beneficial, should not be used as a potential excuse to not deal with other difficulties of special education like under-educated teachers or communicational differences between teacher and student.

We advocate that AI-technologies, in this educational context, should be treated like an assistant rather than a replacement of teachers or other caretakers. Therefore, future research should have a clearer vision of proposed systems as embedded in ”real-life” contexts of schools or homes to ensure that the social aspect of learning with teachers and other students remains, while still finding a way to alleviate emotional stress in these situations.

References

Bhagchandani, Jatin, Manasee Palsule, Sneha Lalwani, Sharmila Sengupta & Jayesh Samtani. 2020. Aid for children with learning disability. *SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology* 12(SUP 1). 129–132. http://smsjournals.com/index.php/SAMRIDDHI/article/view/1920.

Cigale, Matej, Mitja Lustrek, Matjaz Gams, Torsten Kramer, Meike Engelhardt & Peter Zentel. 2018. The improvement of the speech recognition technology (SRT) by artificial intelligence (AI). *Journal of Information Organization* 8(4). 145–151. https://www.dline.info/jio/fulltext/v8n4/jiov8n4_4.pdf.

Committee to Evaluate the Supplemental Security Income Disability Program for Children with Mental Disorders, Board on the Health of Select Populations and Board on Children, Youth and Families, Institute of Medicine and Division of Behavioral and Social Sciences and Education & The National Academies of Sciences, Engineering, and Medicine. 2015. Clinical characteristics of intellectual disabilities. In Thomas F. Boat & Joel T. Wu (eds.), *Mental disorders and disabilities among low-income children*. Washington D.C.: National Academies Press (US). https://www.ncbi.nlm.nih.gov/books/NBK332877/.

Conti, Daniela, Santo Di Nuovo, Serafino Buono & Alessandro Di Nuovo. 2017. Robots in education and care of children with developmental disabilities: a study on acceptance by experienced and future professionals. *International Journal of Social Robotics* 9. 51–62. DOI: 10.1007/s12369-016-0359-6.

European Intellectual Disability Research Network. 2003. *Intellectual disability in Europe: working papers*. http://www.enil.eu/wp-content/uploads/2012/07/Intellectual-Disability-in-Europe.pdf.

Faria, Diego R., Jordan J. Bird, Cintia Daquana, Jhonatan Kobylarz & Pedro P. S. Ayrosa. 2020. Towards AI-based interactive game intervention to monitor concentration levels in children with attention deficit. *International Journal of Information and Education Technology* 10(9). 641–648. http://publications.aston.ac.uk/id/eprint/42048/.

Garg, Shalini & Shipra Sharma. 2020. Impact of artificial intelligence in special need education to promote inclusive pedagogy. *International Journal of Information and Education Technology* 10(7). 523–527. http://www.ijiet.org/vol10/1418-ET004.pdf.

Kharbat, Faten F., A. Alshawabkeh & M. Lynn Woolsey. 2020. Identifying gaps in using artificial intelligence to support students with intellectual disabilities from education and health perspectives. *Aslib Journal of Information Management* 73(1). 101–128. DOI: 10.1108/AJIM-02-2020-0054.

Lepičnik Vodopivec, Jurka & Spela Bagon. 2016. Motivation for using ICT and pupils with learning difficulties. *International Journal of Emerging Technologies in Learning (iJET)* 11(10). 70–75. DOI: 10.3991/ijet.v11i10.5786.

Ouherrou, Nihal, Oussama Elhammoumi, Fatimaezzahra Benmarrakchi & Jamal El Kafi. 2018. Comparative study on emotions analysis from facial expressions in children with and without learning disabilities in virtual learning environment. *Information and Education Technologies* 24. 1777–1792. DOI: 10.1007/s10639-018-09852-5.

Patra, Ganesh & Subhash Chander. 2021. Education of children with disabilities: exploring possibilities with artificial intelligence. *Pedagogy of Learning* 7(3). 29–35. http://pedagogyoflearning.com/wp-content/uploads/4-Ganesh-Patra-Education-of-Children-with-Disabilities-Exploring-Possibilities-with-Artificial-Intelligence.pdf.

Polat, Elif, Tufan Adiguzel & Ozcan Erkan Akgun. 2012. Adaptive web-assisted learning system for students with specific learning disabilities: a needs analysis study. *Educational Sciences: Theory and Practice* 12(4). 3243–3258. https://eric.ed.gov/?id=EJ1003015.

Saad, Swasan, Amal Dandashi, Jihad M. Aljaam & Moataz Saleh. 2015. The multimedia-based learning system improved cognitive skills and motivation of disabled children with a very high rate. *Educational Technology & Society* 18(2). 366–379. http://www.jstor.org/stable/jeductechsoci.18.2.366.

Standen, Penelope J., David J. Brown, Mohammad Taheri, Maria J. Galvez Trigo, Helen Boulton, Andrew Burton, Madeline J. Hallewell, James G. Lathe, Nicholas Shopland, Maria A. Blanco Gonzales, Gosia M. Kwiatkowska, Elena Milli, Stefano Cobello, Annaleda Mazzucato, Marco Traversi & Enrique Hortal. 2020. An evaluation of an adaptive learning system based on multimodal affect recognition for learners with intellectual disabilities. *British Journal of Educational Technology* 51(5). 1748–1765. DOI: 10.1111/bjet.13010.

Tariq, Rabbia & Seemab Latif. 2016. A mobile application to improve learning performance of dyslexic children with writing difficulties. *Journal of Educational Technology & Society* 19(4). 151–166. http://www.jstor.org/stable/jeductechsoci.19.4.151.

Artificial Intelligence in Education

Part I: Didactics and Ethics

2. N. Britt, M. Dammann, N. Niehaus & P. Vorsteher: A human-centered AIED – The ethical implications of AI in the classroom
3. L. Grafweg, L. Witt & P. Heupel: Biases in AI in education
4. K. Breidbach & A. Rönck: Consequences of AI for teachers in K-12 education
5. F. Kock & A. Sommer: AI teacher assistants
6. P. Ondreka, C. Parra Fernandez, F. Wollatz & Q. Zhu: AI-based adaptive gamification in math learning scenarios

Part II: Methods and Technologies

7. M. Duda & N. Dettmer: Comparing different spaced repetition algorithms
8. S. Audichya & E. Korovina: Personalized learning experiences
9. E. Angelescu & J. Näther: Data sets in AI in education
10. N. Daske, R. Abu Ahmad & F. M. Parantean: Automatic lecture transcription
11. K. Sigetova, E. Ganning & T. Obeck: Knowledge representation approaches in adaptive educational AI systems

Part III: Extended Reality and Robots

12. J. Lohkamp, F. Naujoks & J. Moriße: AI and virtual reality in medical education
13. H. Kohnen, A. Ohnesorge, H. Uhlenbrock & D. Witowski: Applications, benefits and drawbacks of 3D holograms in education
14. R. Gratz, L. Lazik, S. Lülf & E. Palme: Humanoid robots in education
15. Y. Heß & M. Rau: AI-based VR Systems for helping learners with special needs

Part IV: Addressing Special Needs

16. K. Vatankhah-Barazandeh & C. Meißner: AI approaches towards reducing the barriers for learners with sensory or physical impairments
17. M. Rehmer & K. Trant: AI-based learning for people in the autism spectrum
18. L. Artman, I. Martynenko & L. Zieba: Using AI for enhancing communication skills and strategies of children with Autism Spectrum Disorder
19. C. Oprotkowitz & I. Grauwelman: AI in education for children with intellectual disabilities