

van der Vorst, Tommy; Jelicic, Nick

## Conference Paper

# Artificial Intelligence in Education: Can AI bring the full potential of personalized learning to education?

30th European Conference of the International Telecommunications Society (ITS): "Towards a Connected and Automated Society", Helsinki, Finland, 16th-19th June, 2019

## Provided in Cooperation with:

International Telecommunications Society (ITS)

*Suggested Citation:* van der Vorst, Tommy; Jelicic, Nick (2019) : Artificial Intelligence in Education: Can AI bring the full potential of personalized learning to education?, 30th European Conference of the International Telecommunications Society (ITS): "Towards a Connected and Automated Society", Helsinki, Finland, 16th-19th June, 2019, International Telecommunications Society (ITS), Calgary

This Version is available at:

<https://hdl.handle.net/10419/205222>

### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

### Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Artificial Intelligence in Education

*Can AI bring the full potential of personalized learning to education?*

**Tommy van der Vorst**  
Dialogic – Innovation & Interaction  
vandervorst@dialogic.nl

**Nick Jelcic**  
Dialogic – Innovation & Interaction  
jelcic@dialogic.nl

## ABSTRACT

In this study we explore the potential impact of educational AI applications in personalized learning. According to Bloom (1984) students that are tutored one-to-one perform two standard deviations better than students who learn via traditional educational methods. Due to the limited amount of teachers and costs associated, personalized one-to-one learning is not generally feasible from a societal point of view. Breakthroughs in the field of machine learning offer promising avenues to aid in personalized learning. AI may hence be the ‘holy grail’ in unlocking the potential of one-to-one learning, by enabling applications to offer personalized teaching to each individual student. We assess the potential impact of AI in personalized learning from a socio-technical perspective. Therefore, we investigate the technological possibilities, as well as any aspects that may impact adoption, e.g. legal, societal and ethical. To conclude we formulate policy options that can stimulate the adoption of AI-driven personalized learning applications.

## INTRODUCTION

The use of technology in education has a long tradition. One tipping point was the introduction of digital technologies in the classroom. This started with the use of desk-top computers in the classroom in the 1980s. Since then, digitisation has permeated all the capillaries of education. Computers were interconnected via global networks, countless teaching methods and content found their way to the student via the Internet, and student tracking systems and other administrative processes were automated. In addition to the computer, more and more other devices were introduced into education that either replace or complement the desktop computer.

Parallel to the development of digital tools and their increased connectivity, the amount of data generated by digital tools and their computing power has grown enormously. This is increasingly leading to the deployment of intelligent systems that can recognise patterns in large amounts of data and that are increasingly capable of imitating human behaviour, especially human reasoning. These systems can therefore perform tasks independently, or support users in performing tasks.

These recent developments in AI might revolutionize education by enabling personalized learning applications. AI system might determine the curriculum and assess the skills level of each individual student. According to Bloom (1984) students that are tutored one-to-one perform two standard deviations better than students who learn via traditional educational methods. Due to the limited amount of teachers and associated costs, personalized one-to-one learning is not

generally feasible from a societal point of view. Furthermore we observe that teachers are experiencing an increased workload. AI powered learning applications could thus not only be beneficial for students but can also increase the productivity of teachers.

In this paper we explore the impact of AI in education from a socio-technical perspective. We look at innovations that might enter the market in the next 5 to 8 years. First, we examine the underlying technologies that enable AI, especially *machine learning*. We look at current applications of AI in education. We conclude this part by giving an overview of properties of AI systems in education. Next, we determine what the impact of these AI systems from a societal, legal and ethical point of view. Finally, we identify the top 5 opportunities and threats related to the use of AI in education and give policy recommendations.

## METHODOLOGY

A multi-method approach was used to answer the research question. The exploration of the applications and use of AI in education is largely based on literature studies and interviews with various stakeholders within and outside the educational field. In addition, an expert workshop was organised to determine the final set of opportunities and risks. One limitation of this study is that it concerns a specific context, namely the Dutch educational system. Nevertheless, many of the findings can be applied to education systems in a more general sense. The AI applications and ethical considerations are particularly relevant to education in general. The legal, and to a lesser extent the recommendations for policy, are more context-specific and will be particularly relevant for the Netherlands and countries with a similar legal and educational system.

### *Interviews*

The set of interviewees included academics in the field of AI, educational science and educational innovation in the Netherlands. In addition, school leaders, teachers, and developers of AI applications in education were interviewed in order to create an encompassing picture of the possibilities of AI in education. Finally, ethicists were consulted in order to reflect on ethical dilemmas that arise with the application of AI in education. The selection of interviewees was made partly on the basis of reputation and assumed expertise, and partly on the basis of snowball sampling. Snowball sampling is a method in which participants in the study suggest other participants for follow-up interviews. The advantage of this method is that hard-to-reach populations, such as teachers, still participate in the study via references.

### *Literature review*

The analyses with regard to the international context, technical possibilities and legal aspects are primarily based on (scientific) literature. First of all, as much material as possible was collected around applications of AI in education, after which (partly on the basis of the knowledge gained from interviews) the information was structured and framed. With regard to the analysis of the legal aspects, a systematic approach was used, starting with a legal interpretation of AI in education. This is then translated into legally relevant acts and actors, in order to determine which laws and regulations apply to them. This assessment leads to a discussion of the legal implications per action and regulation.

## HISTORY OF AI

The history of AI goes further back than many would suspect: the dream of automating human behaviour and reasoning can be traced back to antiquity. In Hellenistic myths we see bronze automatons such as Talos, which protected Crete from pirates. In the Middle Ages, AI ideas can also be found in the stories about the Golem of Prague. Modern ideas about AI go back to the rise of the computer with key figures such as Turing, Walters and Minsky.

The term "artificial intelligence" was introduced in 1956 by scientist John McCarthy. The scientific field has since experienced a number of cycles; highs in which AI was strongly hyped, followed by disappointment and criticism. The AI field has had three such major waves so far. The inception took place around the 1950s and 1960s, around the period that the term was created, driven by pioneers at MIT and Stanford. In the 1970s, however, the research budgets were cut considerably, because AI was unable in practice to translate (for example) texts from Russian into English - an application for which there was a demand at the time (due to the Cold War) and AI was expected to be able to fulfil. The classic example is the poor translation from "The mind is willing, but the flesh is weak" into Russian, to "the vodka is good but the meat is spoiled" into English, by an AI.

In the eighties we see Japan strongly committed to AI to help its industry move forward. The United States and the United Kingdom followed quickly. During this period, AI was dominated by expert systems. Expert systems emulate certain specific tasks/actions of people. In these systems the 'intelligence' is completely programmed by hand. Therefore, the AI could not learn new tasks without a person programming the new rules that are necessary. The systems can best be described as a pre-programmed 'decision tree' that systematically assesses if/else statements to derive a decision. Because experts systems used hand crafted rules, it turned out that the expert systems did not work as well as was expected. It was often too difficult for the AI programmer to hand code all rules into an expert system. In the 1990s, therefore, the attention for AI once again declined sharply.

Recently, AI has had a new revival with pioneers such as Andrew Ng, Geoffrey Hinton and Yann LeCun. They laid the foundation of *deep learning*. Deep learning is a special class of *machine learning*. Machine learning is a set of algorithms that learns from data. Challenges that had previously been considered impossible suddenly turned out to be feasible with deep learning. For example, the AI developed by Google named AlphaGo, which defeated the world champion Lee Sedol in the Go game in 2016. Until then, it was thought that only with human intelligence (and intuition) the Go game could be played at a high level. The Go game has  $10^{172}$  possible moves – numerous magnitudes higher than chess. These are too much moves for traditional methods (e.g. with a minimax tree search) to determine the most optimal strategy.

In deep learning, a mapping between the input and output of a model is iteratively searched for in the form of a series of mathematical transformations. These transformations are loosely inspired by the way our human brain works: a so-called 'neural network'. In our brains, our senses stimulate brain cells - the 'neurons'. Depending on the stimulus, these neurons may or may not send signals to other neurons. Hundreds of billions of neurons together lead to the emergence intelligent behaviour. A rule of thumb that applies to machine learning is that the performance of the model improves as more data is available to train the model. As a result of the ever-

increasing digitisation of processes, a great deal of data is generated that can be used for this purpose. This, combined with ever-increasing computer power and academic breakthroughs, enables the developments in the field of machine learning to grow exponentially.

Deep learning can be seen as a *general purpose technology* and has many applications outside of education that approach or exceed human performance. In the medical world, deep learning is used to diagnose skin cancer and to classify CT scans. In self-driving cars, deep learning is used to determine the direction and speed of the car on the basis of raw camera input. In the fight against fake news, Facebook applies deep learning to assess the authenticity of texts.

## EDUCATIONAL AI APPLICATIONS

There are various types of organisations active in the field educational AI applications. On the one hand, we see that some publishers add a ‘machine learning layer’ to digital educational resources. Here, classification based on hard-crafted rules are being replaced by classification based on machine learning algorithms. In addition, various start-ups are experimenting with applications of AI in education. The major tech giants, such as Google and Facebook, are dominant players in AI's development, but we see a limited presence in the Dutch education market. This is probably because AI requires specific contextual data in education (in this case from Dutch students) that is not easily available.

Academic research in the field of AI is not necessarily fundamental in nature. Practical knowledge is also gained by experimenting with various applications of AI in education. We also see that schools and teachers have started to experiment with AI themselves. IT teachers who have experience with AI apply this to simplify their work. Especially teachers with a business background that have gained experience with machine learning can apply these skills into education.

Generally speaking, AI in education can be used to transfer knowledge and skills, to assess knowledge and skills, to inform instructors of pupils' progress and achievements. Below we give an overview of applications of AI in education as identified (and clustered) during the research.

### *Adaptive learning systems*

An adaptive learning system is a digital learning tool that adapts, as well as possible, to the learner, so that the learning process is optimized, and/or the student performance improve. To achieve this, an adaptive learning system consists of the following components:

1. **Expert model.** This is a model that possesses all the information that will eventually have to be transferred to the student. This model is also aware of the degree of difficulty of the information, and the hierarchical structure between the concepts.
2. **Student model.** This model keeps track of what information the learner has already learned and what skill level the learner has. In fact, it models the way a learner learns.
3. **Instruction model.** This model is specialized in teaching methods. The model knows what the best way to give feedback to a specific student, and when to apply certain didactic strategies such as *productive failure* (first let the student try it himself before

the correct answer is given). The model combines information from the expert model and the student model in order to determine the most suitable instructions.

The above models can be implemented on the basis of relatively simple statistical methods (such as Bayesian statistics) as well as on the basis of advanced deep learning techniques. A paper published by researchers at Stanford University has shown that a *Deep Knowledge Tracing model* is better at modelling the student's knowledge, but (much) less transparent than its Bayesian counterpart (Piech et al., 2015). Something similar generally applies to deep learning-based models compared to more conventional statistical methods: a deep learning model learns a relationship between input and output, but it is not evident what factors contribute to the relationship. With a simpler statistical model, the path from input parameter to output can be traced and understood; deep learning models consists of millions of parameters, and the path can no longer be easily traced and understood.

The three components are eventually brought together in a user interface with which the student interacts. Designing an adaptive learning system is not easy. In an adaptive learning system, there are several goals that need to be optimized simultaneously. Sometimes these goals are conflicting, and sometimes a trade-off between the goals is exists. An adaptive learning system can, for example, aim to teach the student as complex knowledge as possible as quickly as possible. But how does the student experience this? If the questions are too easy, the student does not learn fast enough, but if the questions are too difficult, the student becomes demotivated, which in the long run can lead to a reduction in performance (Cornelisz & van Klaveren, 2018). In addition, an adaptive learning system can 'run wild'. For example, it may happen that a student is generally good at statistics, but has difficulty with the specific subject of regression. A (poorly designed) adaptive learning system could, over time, almost only offer this pupil regression exercises. On the test, however, all subjects are questioned, and the student has not practiced well enough for the other (non-regression) parts.

### ***Automated essay scoring***

Automated essay scoring (AES) is a method of assessing written texts such as papers and essays without the intervention of a human being. AES uses *natural language processing* to assess a written text in terms of content and writing style (Balfour, 2013). In automated essay scoring, a model is trained on the basis of previously, manually assessed texts. Next, machine learning methods are used to determine the extent to which new texts correspond to, or contain elements of, the texts that were used to create the model.

### ***Analytical dashboards***

As the adoption of digital learning resources in education increases, we see that more and more data is generated by these systems. This data can be used to inform the teacher about the performance of students. This is often done in so-called analytical dashboards (West, 2012). These dashboards visualize the progress and performance of students such as pie charts that show how many percent of the students mastering a certain skill, or a line chart that shows the performance of a student over time. Although these descriptive statistic is not necessarily very 'intelligent' in itself, we see an increasing use of predictive modelling (Herodotou, 2017). For example, linking information from a pupil tracking system with personality characteristics to create a predictive model of pupil performance inside and outside the classroom.

### ***Classroom monitoring***

In China we see that AI is not only used for learning and testing, but also that students are monitored by an AI that uses cameras to keep track of which student is paying attention (Connor, 2018). On the basis of this information a student can be addressed, or the teacher can use another method to increase the attention level. Also in European countries we see attention for and research into such technology. In Switzerland and France, for example, we see experiments with technologies that analyse students with cameras in order to provide real-time feedback to the teacher. In Europe we see that classroom monitoring techniques are mainly applied in the context of academic research. In China, the technologies are already being implemented within classrooms. Note that application of the systems in classrooms in China does not seem to improve classroom attention; it can be argued that improvements can be ascribed to effects as described in Jeremy Bentham's Panopticon. A panopticon is a prison in which social control is exercised by making the prisoners feel they are constantly being watched (Norris, 2005). A reaction from a Chinese student speaks volumes: "I dare not be distracted after the cameras have been installed in the classroom. It is as if a pair of mysterious eyes are constantly watching me" (Liang Jun, 2018).

### ***Augmented reality, virtual reality & simulations***

On a global level we see more and more efforts to integrate AI with other digital innovations. In Germany we see research in which AI is integrated with augmented reality by means of smart glasses so that the pupil can immediately see instructions before his or her eyes. German researchers are conducting research into the integration of augmented reality with intelligent tutoring systems, among other things. In Canada, too, research is being carried out into how simulations and virtual reality can be used in combination with AI to enrich the learning experience (Lelei & McCalla, 2018). In these simulation environments, students receive practical instructions from an AI after which they can immediately start working on a problem.

### ***Teachable robots***

A novel application of AI is to use robots to create new didactic methods. For example, AI powered robots are now being developed that need to be taught new information or skills by students (Månsson & Haake 2018). Thanks to the protégé effect, the students learn to understand the lesson material better themselves by explaining it. (Chase et al., 2009) The protégé effect assumes that students learn more quickly when they can explain what they have learned to another student. Evidence for the protégé effect can also be seen outside the classroom: students with younger siblings often have a higher IQ than students who do not (Brjokedala, 2007). A teachable robot can be given 'infinite' explanations, while a person may be inclined to pretend that he understands the explanations in order not to damage his reputation.

## **PROPERTIES OF AI APPLICATIONS**

There is a huge variety of educational AI applications - that follows at least from the examples mentioned above. In order to shape policy with regard to AI-applications in the distinction, it is therefore important to analyse which properties of AI applications are relevant and distinctive. The most important distinguishing properties are listed below.

### ***Didactical versus pedagogical***

Many AI applications in education are aimed at automating didactic activities. These include the selection of teaching material (typically in adaptive learning systems), reviewing tests (such as automated essay scoring), and providing constructive feedback to pupils (intelligent tutoring systems). In addition, instructors and teachers are not only expected to solely transfer knowledge. We also expect students to learn to work together, norms and values and other social competences. An AI can also be used for these purposes. For a group assignment, for example, an AI can make the most diverse selection of students on the basis of the strengths and weaknesses of individual students. (Collaboration.ai , 2018)

### ***Subject specific or cross-curricular***

An AI can be used to transfer knowledge of a particular subject or to learn a specific competence. These subject specific AI applications are already used, for example, to allow a student to practice arithmetic or a new language. A cross-curricular AI application is not only intended to teach one specific competence, but also takes the student's entire performance into account in all subjects. In this way, the AI can not only personalize the content of one specific competence for a student, but also the entire curriculum. Where in primary education the teacher already determines to a certain extent in a cross-curricular way which student is working on what, in higher- and academic education this is less easy. In higher- and academic education we see that it is less easy for teachers to have a good picture of what happens outside their own classroom. In order to compensate for this, teachers in the higher-education are renamed mentors to carry out these cross-curricular tasks. An AI can support this mentoring role by monitoring the performance of a pupil across the subjects and, where necessary, identifying them.

### ***Need for differentiation between students***

In the Dutch education system, the subjects and difficulty are determined by the year and level at which the student finds himself. The student is only admitted to a certain year or level if he or she has achieved a minimum performance over a selection of subjects. For example, at VWO a student is expected to master both mathematics and a foreign language at (at least) the same level of difficulty. When an AI is used in the form of an adaptive learning system, this can make the reduce the practical necessity of differentiating between years and levels. An AI application can make it possible for a student to master one subject at a more advanced level than another. An AI application makes it possible to certify per subject instead of certifying a selection of subjects.

### ***Level of autonomy***

AI applications in education differ widely in the degree of autonomy. There are roughly three levels of autonomy of AI applications in education:

1. The AI provides insight into student performance to the teacher.
2. The AI interprets and provides information on student performance to the teacher.
3. The AI gives direct recommendations to the student (and perhaps parents) without the intervention of the teacher.

When an AI is used for signalling and information purposes, it has a relatively low degree of autonomy. In this situation, the user (the teacher) can decide for him/herself whether to use the information or to ignore it. In this situation, the user can also decide for himself how the



information should be handled. In the case of this informative AI, for example, think of a dashboard with advanced pattern recognition that provides insight into the performance of students. In an adaptive learning system, the AI is in a certain sense already more autonomous. An adaptive learning system can have the autonomy to determine: which student should, at what time, study which course material, and in with which instructions. In the future, even more autonomous systems could be encountered that can act without the intervention of a human being.

In addition, there is also the way in which the AI interacts with users. Choices are made when it comes to providing insight into student performance. How is the information presented, how is the performance of the student made transparent (over time or between students?). In addition, this also presupposes that a teacher has the knowledge and skills to interpret these data correctly. In order to properly assess the data the teacher should know how to interpret probabilities, what a confidence interval is, what is meant by statistical significance, what a type 1 and 2 error is, and so on. There are also various considerations that can be made on how information is *framed*.

### ***Educational styles***

An AI application can be used both to support education in the traditional (often classroom) style and to facilitate new educational styles. Of course, the digital learning systems used today can often be used both inside and outside the classroom. Other applications that fit in with the current way educational style are also conceivable. For example, AI can be used to identify absenteeism within administrative systems. AI can also be used to optimise timetables. One step further towards new educational methods is blended learning. In blended learning, traditional education is combined with digital learning tools. AI is not only used in the digital learning tools but to monitor the classroom with all kinds of sensors, virtual reality and gamification can be used to increase the student's attention and motivation. VR can also be used outside the classroom. VR can also be used to teach more practical skills than, for example, an adaptive learning system. Finally, AI could improve distance learning. With the advent of Massive Online Open Courses (MOOCs), the digital nature of the material means that enormous amounts of data are collected. This data is can serve as input for an AI. MOOCs collect various types of data about students:

1. **Data on the learning process:** while students are studying, all kinds of data are collected, such as: what exercise the student did, how long the student spends on an exercise, and what explanation the student received. These data can be used as input for an adaptive learning system.
2. **Data on the student:** Personal characteristics such as gender, country of origin and age, but also which other MOOCs the student has followed on the platform.
3. **Linking with external sources:** For internet-based services, it is often easier to achieve link with other systems. Many MOOC platforms offer the option of logging in with a Facebook account. Once this has been done, the MOOC platform can enrich existing data with external social media data. In this way, a MOOC can incorporate data that is generated outside the platform.

### ***Accuracy and interpretability***

Generally speaking, AI applications apply advanced predictive models. Furthermore, a prediction is hardly ever 100% accurate (which of course does not mean that a human being would do better). There is often a trade-off between the accuracy of a model and the interpretability of a model. More complex AI-algorithms (especially the recently available deep learning algorithms) are harder to understand, but much more effective. As machine learning is increasingly used to automate human tasks, the necessity of interpretable models grows. Questions such as: what is a model good at, what is a model bad at, and how does a model arrive at a prediction? should be answered when a model is used in practice.

This has resulted in methods being developed to provide insight into complex machine learning models. For example, Shapley Additive Explanation (SHAP) can be used to calculate how a change in input affects the output of a model. In this way, even for the most complex models, it can be demonstrated which properties the model takes into account in a decision and in what way. In addition, the SHAP method can be applied to almost any algorithm, making it a unified method for understanding a model's predictions (Lundberg & Lee, 2017).

In the field of convolutional neural networks, considerable progress has been made in providing insight into these models. With feature visualization, the activations of individual layers of a neural network can be made transparent (Olah et al., 2017). In general, this shows that simple geometric patterns such as edges and corners are recognized in the earlier layers of the network. As we look further into the network, we see more complex combinations of these geometric patterns. Each layer of the network combines the patterns from the previous layers to create more complex features. These insights can be used to determine on what inputs the network pays attention to in creating a prediction.

### **Modularity and portability**

A characteristic of many AI applications is that data is integrated from different sources. This means that providers are inclined to offer total solutions in which the AI application and dataset are integrated. Linking the various systems used in the educational field and exchanging data between them is generally difficult (even outside AI applications).

Initiatives have been launched at the Dutch national level to standardise various aspects of (mainly) digital learning systems. Such standards, for example, facilitate the creation of a single "student identity" between different systems. In the area of AI, we hardly see any such standardisation and a great deal of fragmentation.

## SOCIETAL IMPACT

### *AI and educational institutions*

AI can be categorized as a digital learning resource. In order for AI to be used successfully, similar prerequisites must be met. At the lowest level there must be a good digital infrastructure. The digital infrastructure consists at least of a broadband internet connection and devices for students and teachers. Data must also be available to train AI models. This can be data from administrative systems, (meta)data about educational resources, and data generated by the use of (digital) educational resources.

The prerequisites for successful AI applications are not only technical in nature. Teachers, school leaders and administrators must also be able to see the added value of AI applications and use them responsibly. For this it is necessary that staff in the education sector have sufficient digital skills. Next, subject-specific AI applications can be used in the teaching process. Finally, when users have sufficient experience with responsible application of AI, an institution can move towards integration of different AI systems.

### *AI and students*

The impact of AI will differ between the different education sectors. In primary education, AI can lead to a situation where students are less likely to receive a classroom explanation of the subject matter, but are more likely to receive an explanation that is personalized according to their skill level and learning style. This will lead to students spending more time in the classroom in front of a computer (or at least a screen). Nevertheless, we expect that there will still be a lot of personal contact (needed) between student and teacher.

In secondary education, the impact of AI for the student will be less than in the primary education. AI applications will mainly be found in the educational methods that are already used independently. At the moment (and in the near future), AI systems are particularly suitable for learning and testing factual knowledge. Due to the discrete nature of STEM subjects, the impact of AI will be greater for these subjects than for 'alpha' subjects. An AI is better at assessing whether a mathematical problem has been correctly solved than to assess the arguments in an essay.

In higher education, the impact of MOOCs may be very large. When MOOCs start integrating AI, it is not inconceivable that the same knowledge and skills can be learned via a MOOC but: faster, at lower costs, and from a more prestigious educational institution than in traditional higher education. If this is the case, physical presence in a classroom will no longer be necessary, which will make it easier for students to follow the classes organized by international institutions.

### *AI and teachers*

In the Dutch education system, the teacher plays a central role. The teacher is free to decide how he or she organises the lesson. The design of an AI must therefore be approached from a *design thinking perspective*. In design thinking, the needs of the user are central in the development of a technology. In order to successfully implement AI in education, the AI must be flexible in its use and must fit with the different learning styles of teachers.

The greatest impact of AI, in addition to the users of the applications (pupils and students), will be experienced by instructors. Some teaching tasks can be automated by AI, some of the existing tasks will become more important, and there will also be new tasks introduced by applying AI in education. Among the tasks that can be taken over, we mainly see the selection of course material and assessment/grading. The provision of feedback to the student can also be taken over by an AI. Because some of the tasks can be automated by an AI, this leaves room for the teacher to pay more attention to tasks that are difficult or impossible to automate with AI (e.g. social tasks such as guidance and coaching). The application of AI will also result in new tasks. In order to be able to apply an AI responsibly, an teacher must be able to interpret the output of an AI and translate it into concrete actions.

### *AI and the educational system*

The impact of AI on the Dutch education system can be potentially large. Within the current education system, students are grouped according to competencies out of practical necessity. This makes it easier for teachers to teach a subject matter at a certain difficulty level. When an AI is used for personalised learning, the emphasis will shift from generic 'education levels' to an emphasis on subjects.

AI can eliminate the need for standardised tests. One criticism of standardised tests is that it is a snapshot, and is not always a good representation of a student's performance. With AI the performance of a learner can be monitored throughout the learning process. This eliminates the need for testing at specific moments and creates a better picture of a student's abilities.

## **LEGAL IMPACT**

Under Dutch law, only persons can have rights and/or obligations. These legal subjects can be divided into natural persons (a flesh-and-blood person) and legal entities (such as a BV, a foundation (such as some schools) or the State (such as the Ministry of Education, Culture and Science)). Can we include AI among one or the other?

One could argue that an AI has some characteristics of a person: it is able to learn and make decisions independently of humans. On the other hand, AI is certainly not flesh and blood, and according to the Dutch law system, AI cannot be considered a legal entity, since it is not provisioned for. The conclusion is therefore that AI infrastructure is not a natural or legal person, and therefore does not have independent rights and obligations.

This implies that behind 'the actions of an AI' there is always a person to whom these actions can be assigned, and who carries the related rights (e.g. intellectual property rights) and/or obligations (e.g. liability for AI-based decisions from which damage results). This can be done in accordance with the statutory regulations that apply to aid cases and assistants, and legal representatives.

After reviewing Dutch laws, it can be concluded that the legal status of an AI in education will always depend on the specific application. Nevertheless, the following can be said about this in a general sense:

1. The application of AI in education touches on various generic regulations of various kinds. Some of these regulations are of supplementary law, which means that they can easily and often be deviated from by contract, such as the regulations of the Copyright Act and the Database Law.
2. The flexibility of these private laws is much less evident in the rules of public law, such as the rules on disclosure and re-use of administrative or public sector information. The applicability of these regulations may discourage developers from entering into partnerships with institutions if this would mean that their knowledge would become obsolete. Supplementary contracts on intellectual property rights can (partly) offer a helping hand here. However, it has to be ensured that cooperation with market parties does not ultimately lead to the creation of dominant positions that may conflict with rules on competition.
3. Decisions taken by institutions are, of course, entirely governed by the general rules of administrative law, including the general principles of sound administration. Applying these principles to AI can be difficult because they inherently clash with the black box that AI creates. But it does not seem to be an unsolvable problem.
4. Rules on the protection of personal data (GDPR) are highly applicable to AI. The nature of the personal data in education and the character of AI very difficult to imagine that large-scale AI would be permitted without a statutory regulation. Of course, this will change completely if the data to be used are no longer personal data.
5. If something goes wrong due to the application of AI, there do not seem to be any noteworthy problems with the application of the laws on liability. This set of laws is very open and flexible so that those who are disadvantaged by AI applications should not be put at risk.
6. The education-specific sectoral rules in themselves leave a great deal of freedom and do not seem to contain any harsh restrictions. However, it will always be necessary to act with regards to the principles of education law. In addition, the interests of the child will always have to form a central and ethical point of reference, and the question will always arise as to whether this is the case with AI.

## ETHICAL IMPACT

Regardless of whether certain AI applications are technologically *possible* or legally *allowed*, it is important to determine whether certain applications are *desirable* and *wanted*. Within education, there is the perception that this ethical judgement ultimately lies with teachers and schools. Below, we will discuss a number of ethical aspects that are specifically relevant to the application of AI in education. Norms and values reflect the wishes of a society. But not all incorporated in the law. Norms and values not only apply use of a technology but can also be (consciously or unconsciously) incorporated into the technological design. For example, an engineer may have certain uses in mind when designing (in line with the engineer's norms and values) that are not necessarily in line with the user's.

In education, various stakeholders are involved (primarily the teachers/institution, pupil and parents), each with their own norms and values. An AI application also contains (implicit) norms and values that do not necessarily correspond to the norms and values of the various parties involved.

In this study, various ethical considerations have been identified that are relevant to the use of AI in education. On the basis of this set, those involved can form an opinion about AI applications by assessing the considerations within their own standards and value framework. Note that many of the considerations mentioned are not new and are also relevant in a more general sense (either for AI or for education in general).

### ***Transparency, explainability, and predictability***

When an AI is used to support decision making processes or makes fully autonomous decisions, the user delegates part of the ethical considerations to an AI. When a driver of a self-driving car hands over control to the car, he is not excused from the moral consequences. It is therefore important that the user should be aware of the possible outcomes, and is able to assess whether these fit within his ethical framework. The following matters must be known about an AI application:

- On what data was the AI trained, and is it possible that certain biases may occur in it (which may lead to a bias in decisions or conclusions of the AI)?
- To what extent does the AI continue to learn, and how is it ensured that no bias arises?
- How is the underlying algorithm of the AI structured, and how does the AI deal with examples it has never seen before? Under what circumstances are outcomes of the algorithm correct or incorrect?

### ***Unbriable***

An AI must (as far as possible) make objective choices and may not be susceptible to outside influences. If an AI can be manipulated, it can be exploited. One of the most common criticisms of AES systems is that they reward students for certain word usage rather than content. For example, an AES system can be fooled by using language that is associated with complex reasoning and not focusing on content. Think of excessive use of words such as "nevertheless" instead of "but".

### ***Privacy and self-determination***

The use of AI applications can be in conflict with privacy and self-determination of those involved - the (partial) modelling of a student can be seen as a violation of this, especially if no restrictions are placed on the possible use of such a model.

### ***AI in different ethical theories***

The above aspects can be approached from different ethical frameworks. From a Kantian perspective, rules are determined in advance, which an AI application may not violate under any circumstances (i.e. the end never justifies the means). In a utilitarian perspective, the moral status is determined by the outcome: if an unexplainable AI does demonstrably outperform alternatives, it would probably be unethical not to use the AI. Finally, virtue ethics argues on the basis of the intrinsic properties of the application, which in this context means that the design principles of the AI that must be subject to ethical evaluation. In addition, goals are central in virtue ethics (teleology). From this perspective, one can speak of ethical use of AI when it is used for virtues purposes. Below, we will examine in more detail how these ethical frameworks can be applied to an adaptive learning system that makes decisions that are unexplainable.

When an AI is integrated with an adaptive learning system, from a Kantian perspective, it cannot be called ethical use if the underlying algorithms are unexplainable. An AI application, however good it may be, will always have to comply with the ethical principles. Even if the learner learns significantly faster in the adaptive environment, its use is not ethically justified because it is not explainable how the learning material is personalized.

In a utilitarian perspective, the pros and cons are weighed up against each other. In the case of the adaptive learning system, it will be necessary to determine whether the improvement in learning outweighs the lack of explainability. When there is consensus among all stakeholders that the use of the AI brings more benefits compared to a situation without an AI, the application of AI can be ethically justified.

Goals are central to virtue ethics. In the case of the unexplainable adaptive learning system, one may wonder whether it is being used for a virtues purpose. Does a school use the system to improve the student's performance or to cut back on the teacher population? In the first case one could speak of a valid goal, in the second case one might wonder whether the application of systems with unclear mechanics for cost reduction is a justifiable goal.

## **OPPORTUNITIES AND THREATS**

AI is a technology that can be used in various ways for many different purposes in the educational domain. Below, we consider the greatest opportunities and threats associated with the application of AI in education that have been identified in this study.

### **OPPORTUNITIES**

#### ***Reducing the workload of teachers***

One major opportunity for AI in education is the role that AI can play in solving workload related problems experienced by teachers. In recent years, the teacher has often shown dissatisfaction with the high workload experienced in education. This increased workload is partly due to the additional administrative tasks that teachers have been given with the existing range of tasks as well as due to the reduction of new teachers entering the workforce. AI can support the teacher by automating (administrative) tasks in order to reduce the workload.

Tasks that we expect AI to be able to automate and/or facilitate in the near future are mainly related to proofreading (e.g. highlighting strengths and weaknesses in an essay, after which the instructor primarily assesses these points) and the composition of the course material (with the help of automatic classification of content).

#### ***Personalized learning***

We see that the application of AI in personalized learning is considered a great opportunity. Teachers have limited time and attention and therefore cannot teach each student individually. AI does not have this limitation. This allows an AI to better align education with the wishes of the students. As a result, the teacher is better able to focus his or her attention on 'problem students', and the student goes through the curriculum at his or her own pace and level. The automation of tasks will proceed steadily. We expect that it will start with performing small

tasks such as selecting and practicing course material. Improvements within adaptive learning systems through the implementation of deep learning algorithms could aid in these tasks.

#### ***Supporting the teacher with data-driven insights (learning analytics)***

AI can support the teacher by combining data and making it interpretable. With these *learning analytics*, the teacher can gain holistic and well-founded insights into students. AI can expose cognitive biases and thus make education more fair with regard to ethnicity or gender for example. For example, an AI who is not aware of ethnicity or gender cannot take these variables into account in the school advice. It is known that students with a non-western migration background generally receive lower school advice than native-born students. An AI can correct possible (unconscious) prejudices of a teacher.

#### ***Improved assessment***

In the field of assessment, AI can enable a shift from periodic assessment to continuous assessment. One major criticism of standardized tests is that they are merely snapshots and not a good representation of a student's knowledge. With AI, the knowledge level of a student can be continuously monitored without the necessity of periodic tests.

#### ***Synergy with other digital learning applications***

Finally, AI can increase the effectiveness of existing digital learning resources by replacing a manually programmed rules with rules learned by an AI. In combination with new technologies, such as VR, AR and serious games, synergy can take place through the application of AI. These technologies create a virtual space that can be fully controlled, so that the learning outcomes can be optimised. A possible risk here is that knowledge gained in a virtual environment is generalised to a limited extent to 'the real world', and that other, less formalised knowledge acquired by a student at school (for example on a social level) gets less attention.

## **THREATS**

#### ***Educational objectives can be jeopardised by AI if the emphasis is on technology.***

When the development of AI to focussed data/technology, there is a risk that AI applications will no longer be in line with educational objectives. For example, it is possible that an AI that is optimised for knowledge transfer cannot have incorporated other elements such as motivation. There is also a risk of foreign actors entering the Dutch education market. It is not self-evident that these actors have the same educational goals in mind. As a result, we may lose control of education.

When education is fully automated and students go through the educational process without the intervention of people, one may wonder whether education will not become inhumane in this way. This risk increases when simulations are used to excessively for teaching. There is a chance that knowledge and skills learned in a simulation cannot be applied in the real world.

#### ***Bias in humans in translated to data, and incorporated by AI***

A major risk when deploying AI relates to the data available for training models. In addition to the availability of data, there is also the risk that the data that is available contains human biases. Historical data can contain biases against ethnicity and gender. There is a risk here that AI will



inherit and even reinforce these biases. This can lead to a feedback loop that structurally disadvantages certain minorities.

### *The labour market for teachers*

There is a risk that teachers may need new skills to use AI responsibly. Currently, data-driven education is not part of teacher training, while AI does expect the user to have at least basic digital skills. If teachers cannot keep up with this development, there is a chance that the teacher of the future will be a data scientist with a less strong educational background. However, the risk that 'the teacher' will be automated is very small. If an AI is to be able to take over all the tasks from a teacher, a form of Artificial General Intelligence is required, which we have found to be unrealistic in the near future (Grace et al. 2018).

### *Dependency on black-box models versus the responsibility of teachers*

There is a risk that education will become dependent on AI systems that we do not fully understand. Both teachers and AI need to change in order to overcome this risk. Both the explainability of machine learning models and the knowledge of teachers (in relation to AI) must be increased in order to overcome this risk. A teacher cannot take responsibility for systems that are incomprehensible.

### *Applying AI without the necessary prerequisites*

A number of conditions need to be met before AI can be successfully applied in education. For example, a technical infrastructure and data are required before AI can be applied. If AI is implemented too quickly, there is a chance that the AI will not be able to meet expectations. As a result, AI, just like other innovative learning tools, can lead to great disappointments. This can eventually lead to an aversion to AI in education and may also hinder other (digital) innovations in this domain. In addition, there are a number of legal obstacles (often with good reason) that make it impossible to implement AI strongly overnight.

### *A shift in power*

There are concerns that the introduction of AI will consolidate power with one entity. This entity can be an existing publisher of educational resources or a tech giant. However, recent developments seem to indicate that this risk is manageable. A recent ruling has confirmed that the ownership of data generated in digital educational resources is the property of the school, and not necessarily of the developer of the educational resources. However, it remains a concern for policy makers.

## **POLICY RECOMMENDATIONS**

It seems obvious that the application of AI in education can have a positive impact in various areas. There seems to be little resistance, or even enthusiasm, from the education sector to the use of AI to automate certain (especially administrative and specific educational) tasks. Think of reviewing tests and personalizing a curriculum. At the same time, we see that application of AI with a high impact requires that certain fundamental legal and ethical questions must be answered.

In addition, the role of the teacher with regard to the use of AI is under discussion. On the one hand, teachers can be supported by AI, and are enabled to provide better education. On the other hand, teachers are asked to take responsibility for the decisions of a (unexplainable) AI, to deal with data and assess the value of conclusions from such systems, and to be able to justify the application of AI to those involved.

In order to stimulate applications of AI in education, there are a number of aspects on which policy can be formed and/or on which policymakers (specifically the Ministry of Education, Culture and Science) can take action:

### **1. Stimulate the acceptance of AI in education by teachers, parents and students.**

We propose to ask schools to draw up a vision on the use of AI applications within their own institute. This vision helps to communicate to both our staff and to parents (and indirect students) to what extent the school applies AI, and what the level of ambition is. Of course, schools must be supported in drawing up such a vision. Such an AI vision should at least include the following elements:

- Why does the school apply AI
- For which tasks does the school apply AI (for which tasks *not*)?
- How is the accountability of the AI guaranteed?
- How is it ensured that the AI is not prejudiced?
- How is it ensured that educational goals are not compromised by AI?
- How is the privacy of the student ensured?

### **2. Increase the digital skills of teachers**

Digital skills are a prerequisite for the responsible use of AI. This is necessary for all education sectors (PO/VO/HO). By these digital skills we mean not only handling a computer but also specific data skills that are needed to interpret the output of an AI. Think of a strong statistical basis and general knowledge of some common machine learning algorithms. This means that three actions can be taken:

- More attention should be paid to digital skills on teacher-training courses. Make digital skills an obligation, such as didactic and pedagogic skills.
- The current teacher population must be given further training in order to be able to use AI responsibly.
- Stimulate teachers with a background in data or AI. This can trigger a chain reaction within a school that can promote the adoption of AI.

### **3. Creation of a data infrastructure**

As a rule of thumb, AI applications need a lot of data to perform optimally. The availability of data is therefore a crucial factor in the development of AI applications in Dutch education. The Ministry of Education, Culture and Science may have a role to play in stimulating the provision of data, and in providing frameworks with the purpose of researching and developing AI applications in education. In addition, there is the fear

of bias within data. The Ministry of Education, Culture and Science can help to collect and make public these data sources that can be used to train unbiased AI models.

#### **4. Facilitate experiments with AI in education**

Policy must provide frameworks to the limits of experiments. Availability of data for analysis is one of the most important bottlenecks that could be tackled from a policy perspective. Experimentation also means that not too strict requirements should be set for the results of a pilot project. In addition, funds must also be available to carry out these experiments. Within existing research grants there should be extra room for AI experiments.

#### **5. Stimulate a multidisciplinary approach to the development of AI applications in education**

A multidisciplinary approach is crucial to ensure that educational objectives are not compromised, that user requirements are taken into account, and that the highest technological standards are met. In order to stimulate this multidisciplinary approach, new research grants can be developed that do not focused on merely one domain.

#### **6. Develop a certification for responsible AI use.**

In order to reduce the complexity behind AI for stakeholders such as schools (as purchasers of the applications), parents and students, a certification could be developed, whereby an AI application has to meet certain conditions. For example, a user who is not an AI expert can make a responsible choice when selecting an AI application. Such a certification should at least cover the following aspects:

- Stimulate standardization, data portability, interoperability, in order to ensure that the market for educational AI application functions properly. It should be noted that far-reaching 'top-down' standardisation could actually slow down innovation.
- Develop legal frameworks (in order to creating clarity) and solving legal bottlenecks in the use of AI in education.
- Benchmarks for transparency and explainability of an AI. To what extent must an AI be transparent, which data has been used, which algorithm is used to process it, and which requirements must a explanation of an AI model meet? In practice, the Ministry of Education, Culture and Science can contribute to the establishment of benchmarks that an AI must meet as a minimum requirement. For example, benchmarks for reviewing essays.

## REFERENCES

- Piech, C., Bassen, J., Huang, J., Ganguli, S., Sahami, M., Guibas, L. J., & Sohl-Dickstein, J. (2015). Deep knowledge tracing. In *Advances in neural information processing systems* (pp. 505-513).
- Cornelisz, I., & Van Klaveren, C. (2018). Student engagement with computerized practising: Ability, task value, and difficulty perceptions. *Journal of Computer Assisted Learning*, 34(6), 828-842.
- Norris, C. (2005). From personal to digital: CCTV, the panopticon, and the technological mediation of suspicion and social control. In *Surveillance as social sorting* (pp. 263-295). Routledge.
- Balfour, S. P. (2013). Assessing Writing in MOOCs: Automated Essay Scoring and Calibrated Peer Review™. *Research & Practice in Assessment*, 8, 40-48.
- Chase, C. C., Chin, D. B., Opezzo, M. A., & Schwartz, D. L. (2009). Teachable agents and the protégé effect: Increasing the effort towards learning. *Journal of Science Education and Technology*, 18(4), 334-352.
- Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems* (pp. 4765-4774).
- Olah, C., Mordvintsev, A., & Schubert, L. (2017). Feature visualization. *Distill*, 2(11), e7
- T. Bjerkedala, P.K. G. A. S. J. I. (2007). Intelligence test scores and birth order among young Norwegian men (conscripts) analyzed within and between families
- Grace, K., Salvatier, J., Dafoe, A., Baobao, Z., en Evans, O. (2018). When Will AI Exceed Human Performance? Evidence from AI Experts *Journal of Artificial Intelligence Research*
- Liang Jun, B. (2018). Facial recognition used to analyze students' classroom behaviors *China: People's Daily Online*
- Connor, N. (2018). Chinese school uses facial recognition to monitor student attention in class *Beijing Telegraph*
- Collaboration.ai (2018). Education: shape the minds of the future by adding intelligence to the way students work and collaborate together

Lelei, D., en McCalla, G. (2018). How to Use Simulation in the Design and Evaluation of Learning Environments with Self-directed Longer-Term Learners vol. 10947, Cham: Springer

Månsson, K., en Haake, M. (2018). Teaching Without Learning: Is It OK With Weak AI? vol. 10948, Cham: Springer

West, D. M. (2012). Big data for education: Data mining, data analytics, and web dashboards. *Governance studies at Brookings*, 4(1).

Herodotou, C., Rienties, B., Boroowa, A., Zdrahal, Z., Hlosta, M., & Naydenova, G. (2017, March). Implementing predictive learning analytics on a large scale: the teacher's perspective. In *Proceedings of the seventh international learning analytics & knowledge conference* (pp. 267-271). ACM.