

SÉNAT DE BELGIQUE

SESSION DE 2023-2024

22 NOVEMBRE 2023

Proposition de résolution relative aux mécanismes de sécurité devant encadrer l'utilisation de l'intelligence artificielle dans l'enseignement

(Déposée par M. Rik Daems et consorts)

DÉVELOPPEMENTS

I. INTRODUCTION

«Il permettra à chaque étudiant aux États-Unis, et finalement dans le monde entier, de disposer d'un tuteur personnel de niveau mondial», déclarait Sal Khan en juin 2023, lors d'un test de l'agent conversationnel-tuteur (*chatbot*) expérimental *Khanmigo*, destiné aux écoles et recourant à l'intelligence artificielle (IA).

Sal Khan, le fondateur de la *Khan Academy* et de la *Khan Lab School*, une organisation à but non lucratif, ajoutait qu'il espérait que l'agent conversationnel démocratiserait l'accès des étudiants à des leçons particulières. Il déclarait encore que cet outil pourrait apporter une aide considérable aux enseignants pour des tâches telles que l'organisation des leçons, ce qui leur permettrait de consacrer davantage de temps à leurs élèves (apprenants).

Les compétences linguistiques de l'intelligence artificielle sont telles que certains de ses adeptes suggèrent que les enseignants virtuels seront bientôt capables d'interagir de manière aussi individuelle avec leurs étudiants que les enseignants humains.

«L'intelligence artificielle deviendra aussi capable que les enseignants humains d'offrir une excellente assistance pédagogique», a déclaré Bill Gates lors d'une

BELGISCHE SENAAAT

ZITTING 2023-2024

22 NOVEMBER 2023

Voorstel van resolutie met betrekking tot veiligheidsmechanismen bij het gebruik van artificiële intelligentie in het onderwijs

(Ingediend door de heer Rik Daems c.s.)

TOELICHTING

I. INLEIDING

«Het zal elke student in de Verenigde Staten, en uiteindelijk over de hele wereld, in staat stellen om een persoonlijke tutor van wereldklasse te hebben» verklaarde Sal Khan in juni 2023 in het kader van een experiment met *Khanmigo*, een experimentele *chatbot*-tutor voor scholen die gebruik maakt van artificiële of kunstmatige intelligentie (AI).

Sal Khan, de oprichter van *Khan Academy* en van *Khan Lab School*, een non-profitorganisatie, zei dat hij hoopte dat de *chatbot* de toegang van studenten tot individuele bijlessen zou democratiseren. Hij zei ook dat het leraren enorm zou kunnen helpen met taken als lesplanning, zodat ze meer tijd kunnen besteden aan hun leerlingen.

De taalvaardigheid van kunstmatige intelligentie maakt dat sommige voorstanders suggereren dat virtuele docenten binnenkort in staat zullen zijn om net zo individueel te reageren op studenten als menselijke docenten.

«Kunstmatige intelligentie zal de bekwaamheid verwerven om op eenzelfde niveau als menselijke docenten als uitstekende begeleiders te functioneren» merkte Bill

récente réunion d'investisseurs dans les technologies éducatives. Bill Gates est le cofondateur de Microsoft et un philanthrope (1).

Le recours à l'intelligence artificielle en classe suscite également un enthousiasme de plus en plus grand dans les milieux scientifique et pédagogique:

«Cela est propice à l'apprentissage coopératif, assisté par ordinateur, ainsi qu'à des groupes de discussion asynchrones. Cela permet un apprentissage personnalisé à moindre coût, grâce à un système de navigation basé sur des algorithmes et stimulé par le recours à des évaluations automatisées, à des systèmes de reconnaissance faciale et à des analyses prédictives. Il est donc de plus en plus démontré que l'utilisation de l'IA dans l'enseignement favorise "une transformation des connaissances, des compétences et de la culture" (2).»

Toutefois, l'utilisation de l'IA dans l'enseignement n'est pas dénuée de risques importants, que ce soit pour l'étude proprement dite ou pour l'évaluation. Au Royaume-Uni, on s'est servi d'un algorithme assisté par l'IA durant la pandémie. Cela a débouché sur un véritable scandale en 2020, parce que l'IA semblait évaluer les élèves davantage sur la base de leur localisation qu'en fonction de leurs prestations.

La publication des résultats pour le «niveau A» (*Advanced Level*) a suscité des protestations de la part des élèves. L'algorithme utilisé semblait plus sévère envers les étudiants issus de milieux socioéconomiques plus défavorisés. Il tenait compte à la fois des performances individuelles et de la réputation de l'école. Si une école avait affiché de bons résultats les années précédentes, l'algorithme tablait également sur de bons résultats pour les élèves évalués.

Par conséquent, les élèves moyens fréquentant les «bonnes» écoles obtenaient de meilleurs résultats pour le niveau A que ceux dont les prestations étaient peut-être meilleures mais qui fréquentaient des écoles situées dans des quartiers moins favorisés.

Les critiques portaient surtout sur l'effet inégalitaire de l'algorithme, les résultats des écoles publiques étant sous-évalués et ceux des écoles privées surévalués, ce qui défavorisait les élèves issus de milieux socioéconomiques moins favorisés. Ce biais résultait en partie du comportement de l'algorithme pour de petits groupes

Gates op tijdens een recente bijeenkomst voor investeerders in onderwijstechnologie. Gates is medeoprichter van Microsoft en filantroop (1).

Ook binnen de wetenschap en de pedagogie is men in groeiende mate enthousiast over het gebruik van AI in klaslokalen:

«Het maakt computerondersteund samenwerkend leren mogelijk, evenals asynchrone discussiegroepen. Het staat kostenefficiënt gepersonaliseerd leren toe door middel van een navigatiesysteem gebaseerd op algoritmes, gestimuleerd door het gebruik van geautomatiseerde beoordelingen, gezichtsherkenningssystemen en voorspellende analyses. Hierdoor zijn er groeiende bewijzen voor de rol van AI voor het onderwijs om "een transformatie van kennis, competentie en cultuur te bevorderen" (2).»

Het gebruik van AI in het onderwijs kan echter ook erg grote risico's met zich meebrengen. Niet enkel bij het studeren zelf maar ook bij het beoordelen. In het Verenigd Koninkrijk maakte men tijdens de pandemie gebruik van een door AI geassisteerd algoritme. Dit leidde in 2020 tot een groot schandaal omdat AI leerlingen eerder leek te beoordelen op locatie dan op prestatie.

De bekendmaking van de «A-level» resultaten leidde tot protesten van leerlingen. Het gebruikte algoritme leek strenger te zijn voor studenten met een lagere sociaal-economische achtergrond. Het algoritme hield zowel rekening met individuele prestaties als met de reputatie van de school. Als een school in voorgaande jaren goede resultaten had behaald, voorspelde het algoritme ook goede resultaten voor huidige leerlingen.

Het gevolg was dat gemiddelde leerlingen op 'goede' scholen betere A-level resultaten behaalden dan leerlingen met mogelijk betere cijfers op scholen in minder welvarende gebieden.

De kritiek richtte zich vooral op het ongelijke effect van het algoritme, waarbij resultaten van staatsscholen werden verlaagd en private scholen werden bevoordeeld, wat nadelig uitpakte voor leerlingen met een lagere sociaaleconomische positie. Dit werd deels veroorzaakt door het gedrag van het algoritme bij kleine groepen

(1) <https://www.nytimes.com/2023/06/08/business/khan-ai-gpt-tutoring-bot.html>.

(2) <https://link.springer.com/article/10.1007/s10639-022-11316-w>.

(1) <https://www.nytimes.com/2023/06/08/business/khan-ai-gpt-tutoring-bot.html>.

(2) <https://link.springer.com/article/10.1007/s10639-022-11316-w>.

d'élèves, les écoles privées enregistrant une plus forte croissance des bons résultats que les autres écoles (3).

Un autre risque apparaît lorsque l'IA est aux mains d'un régime autoritaire. La Chine, un pays qui foule déjà aux pieds les droits et libertés des citoyens, tente elle aussi d'introduire l'IA dans les classes. Et elle n'y va pas de main morte.

Selon un récent rapport d'un média contrôlé par l'État chinois, une école secondaire dans l'est de la Chine teste un nouveau système de reconnaissance faciale chargé d'analyser l'implication des élèves en classe.

Ce «système intelligent d'analyse des comportements en classe» scanne le local toutes les trente secondes et enregistre à la fois le comportement des élèves et les expressions de leur visage. Le système est capable d'identifier sept humeurs en analysant le visage d'un élève. Une caméra placée au-dessus du tableau à l'avant de la classe suit par ailleurs six types de comportement: lire, écrire, lever la main, se lever, écouter l'enseignant et s'appuyer sur son bureau (4).

Une chose est sûre: ce système, comme la majorité des applications de l'IA entre des mains chinoises, vise à exercer un contrôle. Il est plus que probable que les résultats et le comportement de l'élève influenceront directement le score qui lui sera attribué en ce qui concerne son crédit social (5). Pareil usage de l'IA pourrait également se rencontrer chez nous.

Il est donc indispensable de mettre en place des 'rails de sécurité' éthiques, capables de nous prémunir contre les dangers de l'intelligence artificielle.

II. QU'EST-CE QUE L'IA?

La technologie de l'intelligence artificielle (IA) se développe à un rythme tel qu'il est difficile de la définir. Le Conseil de l'Europe utilise d'ailleurs deux concepts à cet effet.

L'IA désigne l'ensemble de sciences, théories et techniques dont le but est de reproduire par une machine des capacités cognitives d'un être humain. Les développements actuels visent, par exemple, à pouvoir confier à

(3) <https://www.channel4.com/news/factcheck/factcheck-did-england-exam-system-favour-private-schools>; <https://datanews.knack.be/nieuws/algorithmes-voor-examenuitslag-krijgt-felle-kritiek-in-vk/>.

(4) <https://www.thesun.co.uk/news/22494258/chinese-ai-cam-track-students-concentration-classrooms>.

(5) <https://newatlas.com/china-ai-education-schools-facial-recognition/54786/>.

leerlingen, waardoor private scholen een grotere stijging zagen in hoge cijfers dan andere scholen (3).

Een ander gevaar doet zich voor bij de gevaarlijke mix van AI en autoritaire regimes. China, het land dat sowieso al de rechten en vrijheden van burgers met de voeten treedt, probeert ook AI te introduceren in de klaslokalen. Alleen gaat het hierin heel ver.

Een recent rapport van een door de Chinese Staat gerunde mediabron onthult dat een middelbare school in het oosten van China een nieuw gezichtsherkenningssysteem test dat de betrokkenheid van leerlingen in een klaslokaal in *realtime* analyseert.

Het «intelligente gedragsbeheersysteem voor klaslokalen» scant het lokaal elke 30 seconden en registreert zowel het gedrag van de leerlingen als hun gezichtsuitdrukkingen. Het systeem kan zeven stemmingen identificeren door het gezicht van een leerling te analyseren. Een camera boven op het schoolbord vooraan in het klaslokaal volgt ook zes soorten gedrag: lezen, schrijven, handen opsteken, opstaan, luisteren naar de leerkracht en leunen op het bureau (4).

Voor de goede verstaander is het duidelijk: dit systeem is, net zoals het merendeel van AI in Chinese handen, bedoeld om controle uit te oefenen. De resultaten en het gedrag van de leerling zullen dan naar alle waarschijnlijkheid ook een directe invloed hebben op zijn sociale krediet score (5). AI zou bij ons ook op dezelfde manier kunnen worden gebruikt.

Er is dus nood aan ethische «vangrails» die de gevaren van artificiële intelligentie kunnen opvangen.

II. WAT IS AI?

Artificiële intelligentie (AI) ontwikkelt zich zo snel dat de technologie moeilijk te definiëren is. De Raad van Europa combineert daarom twee definities:

«AI is een verzameling van wetenschappen, theorieën en technieken die het cognitieve vermogen van een mens wil reproduceren door een machine. Nieuwe ontwikkelingen zijn er bijvoorbeeld op gericht om AI

(3) <https://www.channel4.com/news/factcheck/factcheck-did-england-exam-system-favour-private-schools>; <https://datanews.knack.be/nieuws/algorithmes-voor-examenuitslag-krijgt-felle-kritiek-in-vk/>.

(4) <https://www.thesun.co.uk/news/22494258/chinese-ai-cam-track-students-concentration-classrooms/>.

(5) <https://newatlas.com/china-ai-education-schools-facial-recognition/54786/>.

une machine des tâches complexes auparavant réalisées par un humain.

L'IA peut être définie aussi comme l'ensemble des systèmes fondés sur des machines qui, sur la base d'objectifs définis par l'homme, peuvent prendre des décisions, établir des prédictions et formuler des recommandations susceptibles d'avoir une influence sur des environnements réels ou virtuels. Les systèmes d'IA communiquent avec nous et influencent notre environnement de manière directe ou indirecte. Souvent, ils semblent fonctionner de manière indépendante et peuvent adapter leur comportement en fonction de la connaissance qu'ils acquièrent au sujet du contexte.

On peut dire, en d'autres termes, que l'intelligence artificielle repose sur des algorithmes, qui sont une série d'instructions. Les algorithmes sont traités par l'ordinateur en vue de permettre la réalisation de différentes tâches et d'obtenir le résultat escompté. Un algorithme est comparable à une recette de cuisine: c'est une succession d'étapes à suivre pour atteindre le résultat visé, en l'espèce la confection d'un mets déterminé.

Les algorithmes sont les éléments clés de l'IA. L'histoire de l'IA est étroitement liée au développement d'algorithmes sophistiqués et évolutifs. Tous les programmes informatiques sont par essence des suites algorithmiques. Ils sont constitués de milliers de lignes de code qui correspondent à des instructions mathématiques que l'ordinateur doit exécuter pour pouvoir résoudre des problèmes (par exemple, effectuer des calculs numériques, traiter une image et vérifier la grammaire dans un essai). Les algorithmes d'IA sont appliqués dans des domaines qui relèvent par essence du comportement humain, comme la reconnaissance vocale et faciale, l'observation visuelle, la prise de décision et l'éducation. Ainsi, les algorithmes peuvent fournir des instructions pour la quasi-totalité des systèmes d'IA et des applications d'IA que nous pouvons concevoir (6).

Un programme d'IA utilise diverses méthodes d'apprentissage automatique pour obtenir des résultats. L'apprentissage automatique permet aux programmes informatiques d'apprendre et de prendre des décisions de manière autonome, sans qu'une intervention humaine soit nécessaire.

L'apprentissage automatique (*machine learning*) découle des méthodes d'apprentissage statistique et utilise des données et des algorithmes afin de réaliser des tâches traditionnellement effectuées par l'homme.

moeilijke taken te laten uitvoeren die vroeger door de mens werden gedaan.»

«AI verwijst naar op machines gebaseerde systemen die op basis van doelstellingen die de mens definieert beslissingen kunnen nemen en voorspellingen en aanbevelingen kunnen doen die een invloed hebben op reële of virtuele omgevingen. AI-systemen communiceren met ons en beïnvloeden onze omgeving direct of indirect. Vaak lijken ze zelfstandig te werken en kunnen ze hun gedrag aanpassen door te leren over de context.»

Met andere woorden: artificiële of kunstmatige intelligentie omvat een reeks geïnstrueerde handelingen, namelijk een algoritme. Deze algoritmes worden door een computer gevolgd om diverse taken uit te voeren en tot een beoogd resultaat te komen. Het algoritme kan worden vergeleken met een recept, waarbij opeenvolgende stappen worden gevolgd om tot een uitvoerbaar resultaat – in dit geval een gerecht – te komen.

Algoritmes zijn de kernelementen van AI. De geschiedenis van AI is nauw verbonden met de ontwikkeling van geavanceerde en evolutionaire algoritmes. In wezen zijn alle computerprogramma's algoritmes. Ze bestaan uit duizenden coderegels die wiskundige instructies weergeven die de computer volgt om problemen op te lossen (zoals het uitvoeren van numerieke berekeningen, het verwerken van een afbeelding en het controleren van grammatica in een essay). AI-algoritmes worden toegepast op gebieden die we kunnen beschouwen als in essentie menselijk gedrag, zoals spraak- en gezichtsherkenning, visuele waarneming, besluitvorming en onderwijs. Op die manier kunnen algoritmes instructies geven voor bijna elk AI-systeem en elke AI-toepassing die we kunnen bedenken (6).

Een AI-programma hanteert diverse methoden van machinaal leren om resultaten te verkrijgen. Door machinaal leren zijn computerprogramma's in staat om autonoom te leren en tot besluiten te komen, zonder noodzakelijke menselijke tussenkomst.

Machinaal leren (*machine learning*) is afgeleid van statistische leermethoden en maakt gebruik van gegevens en algoritmes om taken uit te voeren die gewoonlijk door mensen worden uitgevoerd. *Machine learning* gaat over

(6) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8455229/>.

(6) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8455229/>.

L'apprentissage automatique revient à laisser à l'ordinateur le soin d'exécuter des tâches sans lui indiquer la marche à suivre.

Les algorithmes d'apprentissage automatique commencent par analyser des données afin d'identifier des modèles, ensuite ils développent leur propre modèle et prévoient, sur la base de celui-ci, ce qu'il pourrait advenir dans le futur.

L'apprentissage automatique peut être considéré, en d'autres termes, comme un processus en trois phases. Le programme commence par analyser et collecter des données, puis il élabore un modèle spécifique permettant la réalisation des tâches prévues et enfin, il entreprend les actions nécessaires et produit les résultats voulus sans aucune intervention humaine. Les applications d'IA les plus connues, telles que les systèmes de recommandation ou de reconnaissance faciale, doivent leur existence à l'apprentissage automatique et à ses principes de fonctionnement (7).

Dans un contexte plus avancé, l'algorithme de l'IA est parfois décrit comme un réseau neuronal, qui présente, sur le plan de la complexité, des similitudes avec les schémas de pensée propres au cerveau humain. Cela traduit la capacité de l'IA à détecter et à interpréter des données, ainsi qu'à agir de manière autonome sur la base de celles-ci (8).

L'intelligence artificielle (IA) est de plus en plus utilisée dans différentes applications comme les moteurs de recherche, les assistants numériques, les agents conversationnels (*chatbots*) et les applications de navigation.

Ces systèmes d'IA puisent des données dans différentes sources et possèdent le potentiel nécessaire pour pouvoir améliorer l'enseignement. Ils permettent de comprendre les besoins d'apprentissage, ils proposent des expériences d'apprentissage personnalisées et aident les écoles dans leur processus décisionnel. Il est capital que les acteurs du monde de l'enseignement et les directions d'écoles soient familiarisés avec l'IA et l'utilisation des données afin de pouvoir appliquer cette technologie de manière efficace et éthique. Dans la proposition de règlement sur l'IA, un système d'IA est défini comme étant un logiciel qui peut, pour un ensemble donné d'objectifs, générer des résultats influençant les environnements avec lesquels il interagit (9).

(7) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8455229/>.

(8) <https://www.vlaanderen.be/kenniscentrum-digisprong/themas/innovatie/artificiele-intelligentie/ai-in-het-onderwijs-waarom-zijn-ethische-richtlijnen-nodig>.

(9) https://learning-corner.learning.europa.eu/learning-materials/use-artificial-intelligence-ai-and-data-teaching-and-learning_fr.

het laten handelen of presteren van computers zonder dat er een regel-voor-regel stap wordt gegeven.

Machine learning algoritmes kijken eerst naar gegevens om patronen te vinden en maken dan een model. Met dat model voorspellen ze wat er in de toekomst zou kunnen gebeuren.

Met andere woorden, *machine learning* kan worden beschouwd als een proces in drie stappen. Eerst analyseert en verzamelt het de gegevens, vervolgens bouwt het een specifiek model voor die verschillende taken en ten slotte onderneemt het de actie en produceert het de gewenste resultaten zonder menselijke tussenkomst. De alom bekende AI-toepassingen zoals aanbevelings-systemen of gezichtsherkenningssystemen zijn allemaal mogelijk gemaakt door de werkingsprincipes van machinaal leren (7).

In een meer geavanceerde context wordt het algoritme van AI soms beschreven als een neurale netwerk, dat qua complexiteit gelijkenissen vertoont met de denkpatronen in menselijke hersenen. Dit verwijst naar het vermogen van AI om gegevens te detecteren, te interpreteren en autonoom te handelen op basis van deze gegevens (8).

Er wordt steeds meer gebruik gemaakt van kunstmatige intelligentie (AI) in verschillende toepassingen, zoals zoekmachines, digitale assistenten, chatbots en navigatie-apps.

Deze AI-systemen verzamelen data uit verschillende bronnen en hebben het potentieel om onderwijs te verbeteren. Ze helpen bij het begrijpen van leerbehoeften, bieden gepersonaliseerde leerervaringen en ondersteunen scholen bij besluitvorming. Begrip van AI en datagebruik is essentieel voor onderwijsactoren en schooldirectie om deze technologie effectief en ethisch te kunnen inzetten. Het voorstel voor een AI-verordening definieert doorgaans een AI-systeem als software die doelgerichte output genereert en interactie heeft met de omgeving (9).

(7) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8455229/>.

(8) <https://www.vlaanderen.be/kenniscentrum-digisprong/themas/innovatie/artificiele-intelligentie/ai-in-het-onderwijs-waarom-zijn-ethische-richtlijnen-nodig>.

(9) https://learning-corner.learning.europa.eu/learning-materials/use-artificial-intelligence-ai-and-data-teaching-and-learning_nl.

A. IA et enseignement

Partout en Europe, les systèmes d'IA sont de plus en plus adoptés dans les salles de classe pour soutenir les pratiques d'enseignement, d'apprentissage et d'évaluation. L'IA a un grand potentiel pour améliorer l'enseignement et pour aider les établissements d'enseignement à perfectionner leur fonctionnement, mais la recherche scientifique relative à l'incidence de l'IA dans l'éducation est encore limitée. Il est donc important de garder une attitude critique.

Les systèmes d'IA peuvent faciliter l'enseignement et l'apprentissage. Ils se subdivisent en trois catégories:

1. «axés sur l'apprenant»;
2. «axés sur l'enseignant»;
3. «axés sur le système».

On distingue quatre cas d'utilisation:

1. *enseignement à l'apprenant*: utilisation de l'IA pour enseigner aux apprenants (axé sur l'apprenant);
2. *soutien à l'apprenant*: utilisation de l'IA pour soutenir l'apprentissage de l'apprenant (axé sur l'apprenant);
3. *soutien à l'enseignant*: utilisation de l'IA pour soutenir l'enseignant (axé sur l'enseignant);
4. *soutien au système*: utilisation de l'IA pour soutenir le diagnostic ou la planification à l'échelle du système (axé sur le système) (10).

B. IA éthique et enseignement

Grâce aux systèmes d'IA, il est possible d'améliorer, d'une part, le transfert des connaissances (matières), l'apprentissage proprement dit et l'évaluation, et, d'autre part, les résultats d'apprentissage et l'efficacité du fonctionnement des écoles. Mais si ces applications d'IA ne sont pas conçues correctement ou si elles ne sont pas utilisées avec précaution, elles peuvent avoir des conséquences néfastes. Les acteurs de l'enseignement doivent être conscients de ce risque et se demander si les systèmes d'IA qu'ils utilisent sont fiables, justes et sûrs, si les données relatives à l'enseignement sont gérées en toute sécurité, si la vie privée des personnes est bien protégée dans ce cadre et si les données en question

(10) https://learning-corner.learning.europa.eu/learning-materials/use-artificial-intelligence-ai-and-data-teaching-and-learning_fr.

A. AI en onderwijs

In Europese klaslokalen worden steeds meer AI-systemen ingezet om diverse onderwijsaspecten te ondersteunen, zoals lesgeven, leren en beoordelen. Hoewel AI veel potentieel heeft om onderwijs te verbeteren en scholen te versterken, ontbreekt het vaak aan empirisch onderbouwd onderzoek naar de daadwerkelijke impact ervan, waardoor kritisch blijven belangrijk is.

AI-systemen kunnen lesgeven en leren vergemakkelijken. Ze worden in drie categorieën onderverdeeld:

1. «leerlinggerichte»,
2. «leerkrachtgerichte»,
3. «systeemgerichte» AI-systemen.

Er zijn vier gebruikssituaties:

1. *leerlingen onderwijzen*: AI om leerlingen te doceren (leerlinggericht);
2. *leerlingen ondersteunen*: AI voor hulp bij leerprocessen van leerlingen (leerlinggericht);
3. *leerkrachten ondersteunen*: AI ter ondersteuning van docenten (leerkrachtgericht);
4. *systemen ondersteunen*: AI voor diagnostiek en systeemplanning (systeemgericht)(10).

B. Ethische AI en onderwijs

Dankzij het gebruik van AI-systemen kunnen leerstofoverdracht, het leren zelf en het beoordelen worden verbeterd, kunnen betere leerresultaten worden verkregen en kunnen scholen efficiënter werken. Als deze AI-toepassingen echter niet naar behoren worden ontworpen of onzorgvuldig worden gebruikt, kan dit schadelijke gevolgen hebben. Onderwijsactoren moeten zich hiervan bewust zijn en zich afvragen of de AI-systemen die zij gebruiken betrouwbaar, fair en veilig zijn en of onderwijsgegevens veilig worden beheerd, of daarbij de privacy van personen wordt beschermd en of die gegevens voor het algemeen belang worden gebruikt. Met de term «ethische AI» wordt bedoeld

(10) https://learning-corner.learning.europa.eu/learning-materials/use-artificial-intelligence-ai-and-data-teaching-and-learning_nl.

sont utilisées dans l'intérêt général. L'expression «IA éthique» vise la mise au point, le déploiement et l'utilisation d'une intelligence artificielle qui garantisse le respect des normes et principes éthiques ainsi que des valeurs fondamentales en lien avec ceux-ci.

Dans le cadre de l'élaboration de lignes directrices en matière d'IA, il y a lieu d'identifier quatre considérations clés qui sous-tendent l'utilisation éthique de l'IA et des données dans l'enseignement, l'apprentissage et l'évaluation: l'action humaine, l'équité, l'humanité et la justification des choix.

L'action humaine renvoie à l'aptitude d'une personne à participer pleinement à la société, les individus déterminant leurs propres choix de vie et étant responsables de leurs actes. Cette capacité d'action sous-tend des concepts tels que l'autonomie, l'autodétermination et la responsabilité.

L'équité signifie que chacun doit être traité équitablement dans l'organisation sociale. Il est nécessaire d'établir des processus clairs pour que tous les utilisateurs aient un accès égal aux possibilités qui leur sont offertes. Il est notamment question ici d'impartialité, d'inclusion, de non-discrimination et de juste répartition des droits et des responsabilités.

L'humanité concerne la considération pour les personnes, leur identité, leur intégrité et leur dignité. Il faut tenir compte du bien-être, de la sécurité, de la cohésion sociale, des contacts significatifs et du respect nécessaires pour garantir un lien humain digne d'intérêt, sans réduire les personnes à de simples objets de données ou à des moyens de parvenir à une fin. L'humanité constitue l'essence même de l'approche de l'IA centrée sur l'humain.

La justification des choix est l'utilisation de connaissances, de faits et de données pour justifier des décisions collectives dans l'environnement scolaire; cette considération exige la transparence et se fonde sur des modèles participatifs et collaboratifs de prise de décision ainsi que sur l'explicité.

Ces considérations éthiques ont une grande valeur intrinsèque et méritent d'occuper une place de choix dans le domaine de l'enseignement. Elles guident les professionnels de l'enseignement et les directions d'établissement dans leurs décisions concernant l'utilisation des systèmes d'IA dans l'enseignement. Ces exigences

op de ontwikkeling, de uitrol en het gebruik van AI waarmee de naleving van ethische normen, ethische beginselen en daarmee verband houdende kernwaarden wordt gewaarborgd.

Bij het opstellen van AI-richtlijnen moeten vier cruciale overwegingen worden geïdentificeerd die de basis vormen voor het ethische gebruik van AI en data in lesgeven, leren en beoordelen: menselijke *agency*, rechtvaardigheid, menselijkheid en gerechtvaardigde keuze.

Menselijke *agency* verwijst naar iemands vermogen om volwaardig deel te nemen aan de maatschappij, waarbij individuen hun eigen levenskeuzes maken en verantwoordelijkheid nemen voor hun acties. Dit ondersteunt principes als zelfstandigheid, zelfbeschikking en verantwoordelijkheid.

Rechtvaardigheid impliceert een eerlijke behandeling binnen de sociale structuur, waarbij duidelijke procedures nodig zijn om gelijke kansen voor alle gebruikers te waarborgen. Dit omvat begrippen als billijkheid, inclusiviteit, non-discriminatie en een evenwichtige verdeling van rechten en verantwoordelijkheden.

Menselijkheid richt zich op het respecteren van individuen, hun identiteit, integriteit en waardigheid. Hierbij moet rekening worden gehouden met welzijn, veiligheid, sociale cohesie, betekenisvolle interacties en respect om betekenisvolle menselijke verbindingen te waarborgen, zonder mensen te reduceren tot louter data-objecten of middelen om doelen te bereiken. Menselijkheid vormt de kern van een mensgerichte benadering van AI.

Gerechtvaardigde keuze betreft het gebruik van kennis, feiten en gegevens om collectieve beslissingen in de onderwijsomgeving te rechtvaardigen, waarbij transparantie en participatieve, samenwerkingsgerichte besluitvormingsmodellen van belang zijn, evenals verklaarbaarheid.

Deze ethische overwegingen hebben intrinsieke waarde en verdienen aandacht binnen het onderwijs. Ze dienen als leidraad voor onderwijsprofessionals en schooldirecties bij hun keuzes omtrent het inzetten van AI-systemen in het onderwijs. Deze ethische voorschriften dragen bij aan het waarborgen van betrouwbare AI-systemen voor

éthiques peuvent contribuer à garantir la fiabilité des systèmes d'IA utilisés dans l'éducation et à faire en sorte que ces systèmes répondent aux préoccupations pertinentes (11).

C. Avantages de l'utilisation de l'IA en classe

Pour les apprenants, l'utilisation de systèmes d'IA est très pratique. Ces systèmes leur permettent d'apprendre à leur manière, car les systèmes comprennent ce que les apprenants maîtrisent et ce qu'ils maîtrisent moins. Ceux-ci peuvent obtenir un *feed-back* immédiat sur leur travail et comprennent ainsi mieux la matière enseignée. L'IA peut également proposer des exercices supplémentaires personnalisés. Les jeux éducatifs et les simulations rendent l'apprentissage plus intéressant et permettent de capter l'attention des apprenants. Si quelqu'un a des besoins spécifiques, l'IA peut adapter les leçons de manière à ce que tout le monde puisse participer. Elle permet aux apprenants de découvrir par eux-mêmes la meilleure façon d'apprendre, ce qui renforce leur confiance en eux et leur motivation.

L'IA permet aux apprenants d'accéder à des sources d'information variées et actualisées, grâce à quoi ils peuvent mieux comprendre le monde qui les entoure et se familiariser avec des technologies de plus en plus présentes dans la société moderne et sur le marché du travail.

Pour les enseignants, l'utilisation de l'IA induit un gain de temps grâce à l'automatisation des tâches routinières telles que l'administration et la paperasserie. Les enseignants peuvent ainsi consacrer davantage de temps à l'interaction directe avec les apprenants et à la création d'outils pédagogiques attrayants. En outre, l'IA collecte et analyse des données sur les performances d'apprentissage, ce qui permet aux enseignants et aux écoles de mieux appréhender les points forts et les points faibles des apprenants.

Il ressort d'une étude de la société de consultance *McKinsey* que les enseignants consacrent 20 à 40 % de leur temps de travail à des activités qui pourraient être automatisées. Cela équivaut à environ treize heures par semaine qui pourraient être consacrées à un meilleur accompagnement des étudiants, avec toutes les conséquences positives qui en découleraient pour les enseignants en termes de satisfaction professionnelle. L'IA peut assister les enseignants dans les tâches de routine. Les processus administratifs peuvent être davantage automatisés et les outils d'apprentissage adaptatif peuvent

éducatieve doeleinden en het aanpakken van relevante vraagstukken (11).

C. Voordelen van het gebruik van AI in de klas

Voor leerlingen is het gebruik van AI-systemen erg handig. Het betekent dat ze op hun eigen manier kunnen leren, omdat de systemen begrijpen wat ze goed en minder goed kunnen. Ze krijgen meteen *feedback* over hun werk en begrijpen de lesstof dus beter. AI kan ook aangepaste extra oefeningen voorstellen. Educatieve spelletjes en simulaties maken leren interessanter en houden de aandacht vast. Als iemand speciale behoeften heeft, kan AI de lessen aanpassen, zodat iedereen kan meedoen. Het zorgt ervoor dat leerlingen zelf kunnen ontdekken hoe ze het beste leren, wat hun zelfvertrouwen en motivatie helpt.

AI stelt leerlingen in staat om toegang te krijgen tot diverse, *up-to-date* informatiebronnen, wat hun begrip van de wereld om hen heen vergroot en hen voorbereidt op het omgaan met technologieën die steeds prominenter worden in de moderne samenleving en op de arbeidsmarkt.

Voor leerkrachten zorgt de inzet van AI voor tijdsbesparing door het automatiseren van routineuze taken zoals administratie en papierwerk. Hierdoor kunnen docenten meer tijd besteden aan directe interactie met leerlingen en het creëren van boeiend leermateriaal. Daarnaast verzamelt en analyseert AI gegevens over leerprestaties, waardoor docenten en scholen inzicht kunnen krijgen in zowel de sterke punten als de verbeterpunten van de leerlingen.

Uit onderzoek van *McKinsey* blijkt dat leerkrachten 20 tot 40 % van hun werktijd spenderen aan activiteiten die kunnen geautomatiseerd worden. Dit komt neer op zo'n dertien uren per week die besteed zouden kunnen worden aan een betere begeleiding van studenten en die een hogere voldoening in het werk met zich meebrengen. Met AI kunnen leerkrachten ondersteund worden in routinematige taken. Administratieve processen kunnen meer worden geautomatiseerd en adaptieve leermiddelen kunnen meer onderwijsstaken overnemen, zoals bijvoorbeeld ondersteuning bij het nakijken van open

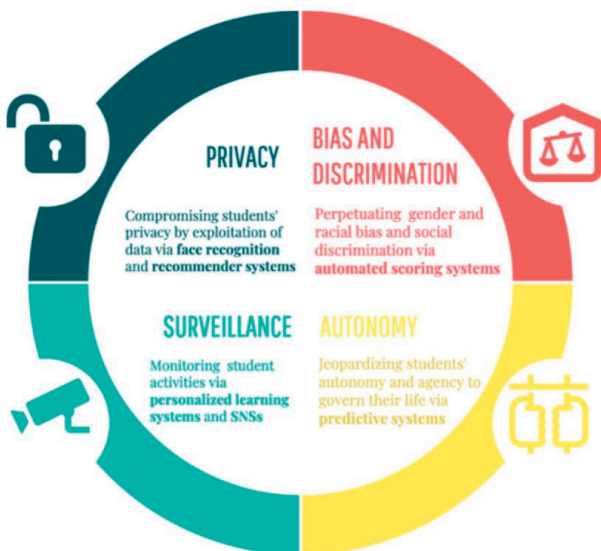
(11) https://learning-corner.learning.europa.eu/learning-materials/use-artificial-intelligence-ai-and-data-teaching-and-learning_fr.

(11) https://learning-corner.learning.europa.eu/learning-materials/use-artificial-intelligence-ai-and-data-teaching-and-learning_nl.

prendre en charge davantage de tâches d'enseignement, telles que l'aide à la correction des exercices ouverts. L'on peut espérer qu'une telle automatisation entraîne une réduction de la charge de travail et renforce l'attrait de la profession.

Les systèmes d'IA peuvent également fournir des instructions pour l'apprentissage des capacités de réflexion, ainsi que proposer et corriger des exercices et donner un feedback. Dans le futur, on verra apparaître des applications d'IA proposant un soutien pour des tâches qui nécessitent des capacités de réflexion plus complexes, comme la réalisation d'analyses. Tel est le cas, par exemple, des systèmes capables de donner un feedback sur des exercices à caractère plus ouvert (12).

D. Inconvénients et dangers de l'utilisation de l'IA



Source: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8455229/>.

Dans l'enseignement, où se joue l'avenir des jeunes, il est capital de prendre divers éléments en considération avant de se fier aux systèmes d'IA.

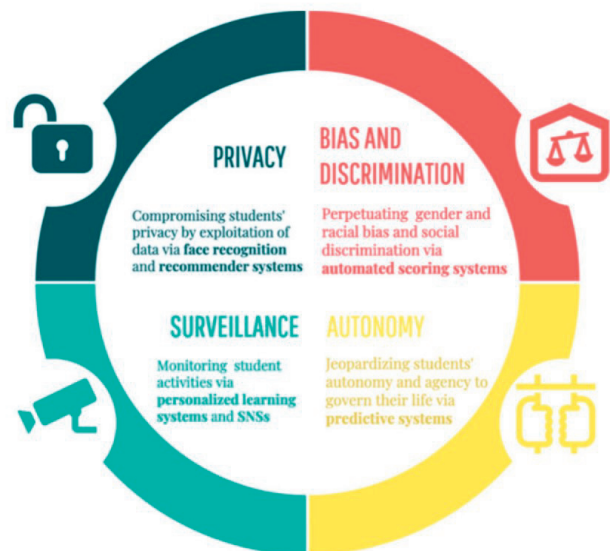
L'absence de responsabilité humaine constitue une préoccupation majeure. Les systèmes d'IA ne sont pas en mesure de remplacer adéquatement l'éthique et la responsabilité humaines, lesquelles sont des éléments fondamentaux dans le paysage de l'enseignement. Si nous voulons préserver l'intégrité et l'éthique du

(12) https://www.serv.be/sites/default/files/documenten/SERV_20210208_Informatierapport_AI_RAP_hoofdstuk%2013.pdf.

opgaven. De hoop en verwachting is dat dit leidt tot minder werkdruk en een aantrekkelijker beroep.

AI-systemen kunnen tevens voor het aanleren van denkvaardigheden de instructie verzorgen en opgaven aanbieden, nakijken en voorzien van *feedback*. In de toekomst komen er AI-toepassingen die ondersteuning bieden bij opdrachten waar complexere denkvaardigheden nodig zijn, zoals het maken van analyses. Dit zijn, bijvoorbeeld, systemen die opdrachten met een meer open karakter kunnen voorzien van *feedback* (12).

D. Nadelen en gevaren verbonden aan het gebruik van AI



Bron: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8455229/>.

Binnen het onderwijs, waar de bouwstenen van de toekomst van jongeren worden gelegd, is het van cruciaal belang om rekening te houden met diverse overwegingen voordat we vertrouwen op AI-systemen.

Het ontbreken van menselijke verantwoordelijkheid vormt een wezenlijke zorg. AI-systemen kunnen niet adequaat de menselijke ethiek en verantwoordelijkheid vervangen, een aspect dat een sleutelrol speelt binnen het onderwijslandschap. We moeten ons realiseren dat de beslissingen die door AI-systemen worden genomen, te

(12) https://www.serv.be/sites/default/files/documenten/SERV_20210208_Informatierapport_AI_RAP_hoofdstuk%2013.pdf.

processus éducatif, nous devons être bien conscients de la nécessité de faire contrôler et évaluer en permanence, par des experts humains, les décisions prises par des systèmes d'IA.

Le récent scandale des allocations familiales aux Pays-Bas illustre les défis liés à l'IA: les systèmes d'IA y ont joué un rôle sans que les parties concernées en soient conscientes.

Le respect de la vie privée et la sécurité sont une autre priorité. L'utilisation de l'IA dans l'enseignement peut éventuellement mettre en péril la protection et la sécurité des données personnelles des étudiants. Elle peut en outre donner lieu à un usage impropre qui n'est plus prioritairement axé sur l'apprentissage mais essentiellement sur le contrôle. Rappelons à cet égard l'exemple de l'utilisation de l'IA dans les classes chinoises. Il est dès lors extrêmement important que les systèmes d'IA répondent à des normes strictes en ce qui concerne le respect de la vie privée et la sécurité.

Le risque de biais et de discrimination apparaît en outre bien réel. Les systèmes d'IA peuvent intégrer des préjugés involontaires, ce qui peut donner lieu à une inégalité de traitement et à de l'injustice dans l'enseignement. Il est essentiel de contrôler et tester régulièrement les systèmes d'IA afin de garantir qu'ils ne sont entachés d'aucun préjugé.

Par ailleurs, un usage trop intensif de l'IA peut réduire l'interaction humaine entre les étudiants et les professeurs, ce qui peut nuire au développement social et à l'intelligence émotionnelle. Des problèmes techniques, comme des bugs et des dérèglements des systèmes d'IA, peuvent entraîner des retards d'apprentissage et des frustrations chez les apprenants et les enseignants. Une dépendance accrue à la technologie est également un effet secondaire possible du recours intensif à l'IA dans l'enseignement. En cas de défaillance de cette technologie, les élèves peuvent éprouver des difficultés à résoudre certains problèmes sans l'aide de la technologie. L'imprécision des systèmes d'IA, résultant du caractère imprécis ou incomplet des données qui les alimentent, peut constituer un défi.

Des problèmes surgissent en outre lorsque de tels systèmes d'IA soulèvent des questions relatives à l'autonomie et, plus spécifiquement, à l'aptitude d'une personne à agir en fonction de ses propres intérêts et valeurs. Les systèmes prédictifs dont le fonctionnement est basé sur des algorithmes mettent en péril l'autonomie des apprenants et des enseignants, ainsi que leur capacité à déterminer eux-mêmes leur comportement. Le recours

à l'IA en tout temps doivent être contrôlés et évalués par des experts humains afin de garantir l'intégrité et l'éthique du processus d'enseignement.

Een recent voorbeeld van de uitdagingen rondom AI is te zien bij de toeslagenaffaire in Nederland, waar AI-systemen een rol speelden zonder dat de werking ervan duidelijk was voor de betrokken partijen.

Privacy en beveiliging zijn nog een ander aandachtspunt. Het gebruik van AI in het onderwijs kan potentieel de privacy en veiligheid van studentengegevens in gevaar brengen. Daarnaast kan het fout geïmplementeerd worden, zodat het zwaartepunt niet meer ligt op het daadwerkelijke leren, maar vooral op het controleren. We verwijzen hiervoor naar het voormelde voorbeeld van China inzake AI. Het is daarom van uiterst belang dat AI-systemen voldoen aan strenge normen op het gebied van privacy en beveiliging.

Verder blijkt het probleem van *bias* en discriminatie een reëel risico te zijn. AI-systemen kunnen onbedoelde vooroordelen vertonen, wat kan resulteren in ongelijke behandeling en onrechtvaardigheid in het onderwijs. Het is essentieel om AI-systemen regelmatig te monitoren en te testen om te waarborgen dat ze geen vooroordeel vertonen.

Daarnaast kan het overmatige gebruik van AI leiden tot een verminderde menselijke interactie tussen studenten en docenten, wat gevolgen kan hebben voor sociale ontwikkeling en emotionele intelligentie. Technische problemen, zoals bugs en storingen in AI-systemen, kunnen leiden tot leerachterstanden en frustratie bij leerlingen en docenten. Een grotere afhankelijkheid van technologie is ook een mogelijk neveneffect van het intensieve gebruik van AI in het onderwijs. Als deze technologie faalt, kunnen er problemen ontstaan bij het oplossen van vraagstukken zonder technologische hulp. De onnauwkeurigheid van AI-systemen kan een uitdaging vormen wanneer de onderliggende gegevens onnauwkeurig of onvolledig zijn.

Er ontstaan ook problemen wanneer dergelijke AI-systemen vragen oproepen die te maken hebben met autonomie, meer specifiek, het vermogen van de persoon om te handelen vanuit haar of zijn eigen belangen en waarden. Voorspellende systemen die worden aangedreven door algoritmes brengen de autonomie van leerlingen en leraren in gevaar en hun vermogen om hun eigen doen te bepalen. Het gebruik van algoritmes om

aux algorithmes pour faire des prédictions sur les actions des individus à partir d'informations les concernant soulève des questions d'intégrité et d'épanouissement personnel. C'est pourquoi il existe également un risque que l'analyse prédictive perpétue les préjugés existants et alimente la discrimination et la stratification sociales.

Il importe par ailleurs de reconnaître que les systèmes d'IA ne peuvent pas remplacer la créativité humaine, qui est un élément fondamental de l'enseignement. Il est donc essentiel de compléter stratégiquement l'IA par de la créativité humaine.

Lors de l'implémentation de systèmes d'IA dans l'enseignement, il faut aussi tenir compte des coûts et de l'accessibilité, de sorte que cette technologie soit disponible pour tous les établissements d'enseignement et toutes les communautés, quelle que soit leur situation financière.

Des intérêts commerciaux majeurs peuvent peser sur les considérations éthiques. Un des principes de base des systèmes éducatifs intelligents est qu'un étudiant interagit avec des interfaces utilisateur adaptatives qui personnalisent les expériences d'apprentissage en tenant compte de l'étudiant et de son niveau actuel d'apprentissage. L'avantage des systèmes d'IA basés sur les données est qu'ils sont capables de traiter des flux de données très complexes en temps réel. Pour les systèmes éducatifs de la prochaine génération, cela implique que ces systèmes ont besoin d'interfaces utilisateur qui collectent en temps réel des données sur le comportement de l'élève (l'apprenant), ainsi que des données historiques pouvant servir à modéliser l'élève («no AI without UI (user interfaces)»).

Il en résulte un intérêt commercial accru pour l'installation de technologies de capteurs et d'interfaces utilisateur dans les classes et l'accès à d'autres sources de données associées à l'apprenant, comme les réseaux sociaux et les plateformes de jeux en ligne (13).

Enfin, une autre préoccupation concerne la perte éventuelle de motivation et d'implication des étudiants qui deviendraient trop dépendants des systèmes d'IA pour réaliser leurs tâches. Il est indispensable que la mise en œuvre des systèmes d'IA continue à stimuler les étudiants à apprendre et à se développer tout en sauvegardant la dimension humaine de l'enseignement (14).

voorspellingen te doen over de acties van individuen op basis van hun informatie roept vragen op over eerlijkheid en persoonlijke ontplooiing. Daarom bestaat ook het risico dat voorspellende analyse bestaande vooroordelen en vooroordelen van sociale discriminatie en stratificatie bestendigen.

Verder is het belangrijk om te erkennen dat AI-systemen geen menselijke creativiteit kunnen vervangen, wat een fundamenteel element is van onderwijs. Het strategisch aanvullen van AI met menselijke creativiteit is daarom van groot belang.

Bij de implementatie van AI-systemen in het onderwijs moeten ook de kosten en toegankelijkheid in overweging worden genomen, zodat deze technologie beschikbaar is voor alle scholen en gemeenschappen, ongeacht hun financiële situatie.

Grote commerciële belangen kunnen mogelijk een impact hebben op ethische afwegingen. Een kernidee van intelligente onderwijssystemen is dat een student interageert met adaptieve *userinterfaces* die leerervaringen personaliseren op basis van de student en zijn actuele leerniveau. Het voordeel van datagebaseerde AI-systemen is dat zij zeer complexe datastromen in *realtime* kunnen verwerken. Voor onderwijssystemen van de volgende generatie impliceert dit dat deze systemen gebruikersinterfaces nodig hebben die *realtime* invoer verzamelen over het gedrag van de leerling als ook historische gegevens, die kunnen worden gebruikt om de leerling te modelleren («no AI without UI (user interfaces)»).

Dit vergroot het commerciële belang voor de installatie van sensortechnologieën en gebruikersinterfaces in het klaslokaal en het verkrijgen van toegang tot andere met de leerling geassocieerde gegevensbronnen zoals sociale media en gameplatformen (13).

Tot slot, een andere zorg is de mogelijke afname van motivatie en betrokkenheid bij studenten wanneer ze te afhankelijk worden van AI-systemen voor hun taken. Het is noodzakelijk dat de implementatie van AI-systemen studenten blijft stimuleren om te leren en zich te ontwikkelen, terwijl we eveneens de menselijke dimensie van het onderwijs intact houden (14).

(13) [https://www.serv.be/sites/default/files/documenten/SERV_20210208_Informatierapport_AI_RAP_hoofdstuk%20\(13\).pdf](https://www.serv.be/sites/default/files/documenten/SERV_20210208_Informatierapport_AI_RAP_hoofdstuk%20(13).pdf).

(14) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8455229/>.

(13) [https://www.serv.be/sites/default/files/documenten/SERV_20210208_Informatierapport_AI_RAP_hoofdstuk%20\(13\).pdf](https://www.serv.be/sites/default/files/documenten/SERV_20210208_Informatierapport_AI_RAP_hoofdstuk%20(13).pdf).

(14) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8455229/>.

E. Le problème de la «boîte noire»

Le problème de la «boîte noire» constitue un défi considérable inhérent à l'intelligence artificielle (IA) et à l'apprentissage automatique (*machine learning* (ML)). Ce phénomène renvoie au manque de transparence et d'interprétabilité des processus décisionnels des modèles complexes d'IA. Lorsqu'un système d'IA fait des prédictions ou prend des décisions, il est parfois difficile de comprendre comme il en est arrivé là. Ce manque de transparence est souvent comparé à une «boîte noire» dans laquelle entrent des données mais dont le fonctionnement et le raisonnement internes conduisant au résultat produit ne sont pas faciles à comprendre.

Le problème de la «boîte noire» présente différents aspects importants:

- la complexité: de nombreux modèles d'IA de pointe, comme les réseaux neuronaux profonds, se composent de multiples couches et de points nœuds reliés entre eux. Cette complexité rend difficilement traçable la manière dont les données entrées sont converties en prédictions exécutables;
- le manque d'explicabilité: les logiciels traditionnels sont conçus sur la base de règles explicites que l'on peut suivre et comprendre. Au contraire, les modèles d'IA et surtout les méthodes d'apprentissage profond (*deep learning*) apprennent des modèles de grands ensembles de données. Leur processus décisionnel est généralement difficilement explicable en langage humain;
- la confiance et la responsabilisation: pour les applications critiques de l'IA, comme dans les soins de santé, les finances et l'enseignement, il est fondamental de comprendre pourquoi un système d'IA a pris une décision déterminée. Un manque de transparence peut éveiller la méfiance et être un frein à la responsabilité en cas d'erreur;
- la partialité et l'intégrité: si les modèles d'IA fonctionnent sur la base de données partiales, il se peut que ces préjugés se reflètent dans leurs prédictions. Le phénomène de la boîte noire qui caractérise les modèles d'IA peut compliquer l'identification et la limitation de tels partis pris.

C'est ce problème qui était en partie à l'origine du scandale évoqué précédemment, provoqué au Royaume-Uni par des algorithmes entachés de préjugés qui évaluaient plus sévèrement certaines catégories d'élèves.

E. Het «*black box*» probleem

Het «*black box*» probleem is een aanzienlijke uitdaging die gepaard gaat met kunstmatige intelligentie (AI) en *machine learning* (ML) systemen. Het verwijst naar het gebrek aan transparantie en interpreteerbaarheid in de besluitvormingsprocessen van complexe AI-modellen. Wanneer een AI-systeem voorspellingen doet of beslissingen neemt, is soms moeilijk te begrijpen hoe het daar toe is gekomen. Dit gebrek aan transparantie wordt vaak vergeleken met een «*black box*» waar *input* ingaat, maar de interne werking en redenering achter de output niet gemakkelijk te begrijpen zijn.

Er zijn verschillende belangrijke aspecten van het black box probleem:

- complexiteit: veel geavanceerde AI-modellen, zoals diepe neurale netwerken, bestaan uit talloze lagen en onderling verbonden knooppunten. Deze complexiteit maakt het moeilijk om te traceren hoe invoerdata wordt omgezet in uitvoerbare voorspellingen;
- gebrek aan verklaarbaarheid: traditionele softwareprogramma's zijn ontworpen met expliciete regels die kunnen worden gevolgd en begrepen. In tegenstelling hiermee leren AI-modellen, vooral diepe leermethoden (*deep learning*), patronen van grote datasets. Hun besluitvormingsproces is doorgaans niet gemakkelijk te verklaren in menselijke termen;
- vertrouwen en verantwoording: bij kritieke AI-toepassingen zoals gezondheidszorg, financiën en onderwijs is het cruciaal om te begrijpen waarom een AI-systeem een bepaalde beslissing heeft genomen. Een gebrek aan transparantie kan leiden tot wantrouwen en het belemmeren van verantwoordingsplicht wanneer dingen fout gaan;
- vooringenomenheid en eerlijkheid: als AI-modellen worden getraind op basis van vooringenomen gegevens, kunnen ze deze vooroordelen mogelijk doorgeven in hun voorspellingen. Het *black box*-karakter van de modellen kan het moeilijk maken om dergelijke vooroordelen te identificeren en te beperken.

Dit probleem lag ten dele aan de grondslag van het eerder besproken schandaal in het Verenigd Koninkrijk met oneerlijke algoritmes die bepaalde categorieën van leerlingen strenger beoordeelden.

Lors de la formation du gouvernement des Pays-Bas en 2021, il avait été décidé d'inscrire dans l'accord de gouvernement que le pays s'emploierait à développer la numérisation. La coalition promettait également de mettre en place une autorité de contrôle des algorithmes. Une disposition légale prévoit ainsi que le contrôle des algorithmes porte sur la transparence, la discrimination et l'arbitraire (15).

Cette décision n'était pas le fruit du hasard: l'affaire des allocations familiales aux Pays-Bas avait montré que le fisc faisait confiance à des algorithmes qu'il comprenait à peine lui-même et donc qu'il pouvait à peine expliquer. Dans l'enseignement, il est capital de reconnaître que le fait de s'en remettre totalement aux décisions de l'IA peut non seulement s'avérer inéquitable, mais aussi avoir des conséquences qu'il convient d'évaluer avec le plus grand soin.

Les jeunes se trouvent à un stade de leur développement où interviennent leur croissance personnelle, leurs talents et leurs aspirations. Les décisions relatives à leur filière d'enseignement, à leurs études et à leurs performances universitaires peuvent avoir des répercussions à long terme sur leur futur parcours. Bien que les systèmes d'IA soient un gage d'efficacité et soient très utiles pour l'analyse des données, il ne faut pas oublier que l'IA ne sera jamais en mesure de remplacer des aspects humains tels que l'empathie, la créativité et la compréhension.

III. LA NÉCESSITÉ D'INSTAURER DES MÉCANISMES DE SÉCURITÉ POUR ENCADRER LE RECOURS À L'IA DANS LES ENVIRONNEMENTS D'APPRENTISSAGE

La mise en œuvre de systèmes d'IA contribue à améliorer les méthodes d'enseignement, les processus d'apprentissage et les procédures d'évaluation, ce qui permet *in fine* d'améliorer les résultats éducatifs et l'efficacité opérationnelle des établissements d'enseignement.

Toutefois, si ces applications d'IA sont mal conçues ou ne sont pas utilisées avec précaution, elles peuvent avoir des conséquences néfastes. Les professionnels de l'enseignement doivent être conscients de ces dangers et évaluer si les systèmes d'IA qu'ils utilisent sont fiables, équitables, sûrs et éthiques.

La gestion des données relatives à l'éducation doit se faire de manière sécurisée et respectueuse de la vie

Bij de regeringsvorming in Nederland in 2021 werd in het coalitieakkoord beslist dat het land zich inzet voor meer digitalisering. Daarin werd ook beloofd dat een algoritmetoezichthouder zal aangesteld worden. Er wordt ook wettelijk geregeld «dat algoritmes worden gecontroleerd op transparantie, discriminatie en willekeur»(15).

Deze beslissing was geen toeval: de Nederlandse toeslagenaffaire liet zien dat de Belastingdienst vertrouwd op algoritmes die ze zelf amper begreep, laat staan kon uitleggen. In het onderwijs is het van cruciaal belang om te erkennen dat het uitsluitend overlaten van beslissingen aan AI niet alleen onrechtvaardig kan zijn, maar ook potentiële gevolgen heeft die we zorgvuldig moeten afwegen.

Jonge individuen ondergaan een fase van ontwikkeling die verweven is met hun persoonlijke groei, talenten en aspiraties. Beslissingen met betrekking tot onderwijsrichting, studierichtingen en academische prestaties kunnen een langdurige invloed hebben op hun toekomstige trajecten. Hoewel AI-systemen zeker bijdragen aan efficiëntie en gegevensanalyse, moeten we in acht nemen dat menselijke aspecten zoals empathie, creativiteit en begrip niet zomaar door AI kunnen worden vervangen.

III. DE NOOD VOOR VEILIGHEIDSMECHANISMEN OM TRENTA IN LEEROMGEVINGEN

De implementatie van AI-systemen draagt bij aan de verbetering van onderwijsmethoden, leerprocessen en evaluatieprocedures, wat leidt tot verbeterde educatieve resultaten en verhoogde operationele efficiëntie van scholen.

Echter, wanneer deze AI-toepassingen ontoereikend worden ontworpen of onzorgvuldig worden ingezet, kunnen er nadelige gevolgen optreden. Onderwijsprofessionals dienen zich van deze potentiële risico's bewust te zijn en te overwegen of de AI-systemen die zij inzetten betrouwbaar, rechtvaardig, veilig en ethisch verantwoord zijn.

Het beheer van onderwijsgegevens moet veilig verlopen, met inachtneming van de privacy van individuen

(15) https://www.parlement.com/id/vloreou0m8t6/regeerakkoord_2021.

(15) https://www.parlement.com/id/vloreou0m8t6/regeerakkoord_2021.

privée des personnes et les données en question doivent contribuer au bien-être général (16).

Les établissements d'enseignement sont généralement chargés de traiter des quantités importantes de données relatives à l'éducation, en ce compris des informations à caractère personnel sur les apprenants, les parents, le personnel, la direction et les fournisseurs. Les données collectées, utilisées et analysées dans le domaine de l'enseignement sont souvent appelées «données éducatives».

Il s'agit de données enregistrées dans les systèmes d'information concernant les apprenants, par exemple les performances scolaires ou académiques, les informations parentales, les notes d'évaluation et des données détaillées résultant de l'utilisation d'outils numériques. Lorsque les apprenants interagissent avec des dispositifs numériques, ils génèrent des traces numériques.

De même, l'utilisation d'environnements informatiques pour l'apprentissage humain en salle de classe, pour apprendre les mathématiques ou les langues vivantes par exemple, produit des traces de l'activité d'apprentissage. Les données collectées dans ce cadre peuvent être combinées pour cerner le comportement en ligne de chaque apprenant. De telles données de traçage, qui fournissent des informations sur l'utilisation du numérique et les activités d'apprentissage, sont souvent utilisées pour l'analyse du processus d'apprentissage. Les données contenues dans les systèmes d'information sur les apprenants peuvent aussi être utilisées pour planifier les ressources d'apprentissage et les cours, ainsi que pour établir des prévisions concernant les taux d'abandon et les besoins d'accompagnement.

Compte tenu des grandes quantités de données nécessaires à l'entraînement des systèmes d'IA, de la nature automatique des algorithmes et de la possibilité de les déployer à grande échelle, l'utilisation de l'IA soulève d'importantes questions relatives aux données à caractère personnel, à la protection des données et à la vie privée. Les établissements d'enseignement sont tenus de s'assurer que toutes les données qu'ils traitent sont stockées de manière confidentielle et sécurisée. Il est en outre crucial qu'ils mettent en œuvre des politiques et procédures appropriées pour garantir à la fois la protection des données à caractère personnel et leur utilisation éthique, dans le respect du règlement général sur la protection des données (RGPD) (17).

(16) <https://www.vlaanderen.be/kenniscentrum-digisprong/themas/innovatie/artificiele-intelligentie/ai-in-het-onderwijs-waarom-zijn-ethische-richtlijnen-nodig>.

(17) https://learning-corner.learning.europa.eu/learning-materials/use-artificial-intelligence-ai-and-data-teaching-and-learning_fr.

en de gegevens dienen bij te dragen aan het algemeen welzijn (16).

Scholen zijn doorgaans belast met het verwerken van aanzienlijke hoeveelheden onderwijsgerelateerde gegevens, waaronder persoonlijke informatie van leerlingen, ouders, personeel, directieleden en leveranciers. De informatie die binnen het onderwijs wordt vergaard, aangewend en geanalyseerd, wordt vaak omschreven als «onderwijsgegevens».

Deze omvatten gegevens die geregistreerd zijn in informatiesystemen voor leerlingen, zoals academische prestaties, ouderlijke informatie, cijfers en gedetailleerde gegevens die voortkomen uit het gebruik van digitale hulpmiddelen. Terwijl leerlingen digitale apparaten hanteren, laten ze digitale sporen achter.

Op een analoge wijze genereren leeractiviteiten sporen wanneer studenten bijvoorbeeld wiskunde of vreemde talen leren met behulp van intelligente tutoringsystemen binnen de klasomgeving. Deze verzamelde data kunnen gecombineerd worden om het online gedrag van elke leerling vast te leggen. Dergelijke traceergegevens, die wijzen op digitaal gebruik en leeractiviteiten, worden vaak ingezet voor analyses van het leerproces. Informatie opgenomen in leerlinginformatiesystemen kan tevens dienen voor de planning van leermiddelen en lessen, en voor het voorspellen van uitvalpercentages en begeleidingsbehoeften.

Aangezien het trainen van AI-systemen aanzienlijke hoeveelheden data vereist, gecombineerd met de automatiserende aard van algoritmes en de mogelijkheid om deze op grote schaal in te zetten, roept het gebruik van AI belangrijke vraagstukken op met betrekking tot persoonlijke data, gegevensbescherming en privacy. Scholen dienen ervoor te zorgen dat alle gegevens die zij verwerken, betrouwbaar en veilig worden opgeslagen. Daarnaast is het essentieel dat zij passend beleid en geschikte procedures implementeren om zowel de bescherming als het ethisch gebruik van persoonlijke data te waarborgen, dit alles in overeenstemming met de algemene verordening gegevensbescherming (AVG) (17).

(16) <https://www.vlaanderen.be/kenniscentrum-digisprong/themas/innovatie/artificiele-intelligentie/ai-in-het-onderwijs-waarom-zijn-ethische-richtlijnen-nodig>.

(17) https://learning-corner.learning.europa.eu/learning-materials/use-artificial-intelligence-ai-and-data-teaching-and-learning_nl.

IV. LIGNES DIRECTRICES ÉTHIQUES SUR L'UTILISATION DE L'INTELLIGENCE ARTIFICIELLE (IA)

A. Au niveau européen

La Commission européenne a déjà pris plusieurs initiatives pour promouvoir, à différents niveaux, la recherche sur l'IA et la mise en œuvre de celle-ci, comme en témoignent les initiatives suivantes (18).

Afin d'unir les forces et de créer des synergies, la Commission européenne a collaboré avec vingt-cinq États membres pour élaborer une déclaration commune sur l'IA.

Elle a en outre créé l'Alliance européenne pour l'intelligence artificielle, une plateforme qui a pour mission de rassembler et de partager des expertises, connaissances et ressources en vue de promouvoir le développement et la mise en œuvre de l'IA en Europe.

La Commission européenne a également publié une importante communication sur l'intelligence artificielle. Cette communication présente, dans les grandes lignes, la vision stratégique de l'Union européenne en matière d'IA et souligne le rôle crucial que celle-ci aura dans l'avenir de l'Europe.

Pour traduire cette vision en actes, la Commission européenne a élaboré un plan coordonné sur l'intelligence artificielle, qui jette les bases d'efforts conjoints visant à promouvoir le développement, la mise en œuvre et la régulation des technologies de l'IA en Europe.

Dans le cadre de son engagement en faveur d'une utilisation éthique et responsable de l'IA, la Commission européenne a également élaboré un guide de lignes directrices éthiques. Ce guide contient des lignes directrices et principes à suivre dans le cadre du développement et de l'application de l'IA, afin de préserver des valeurs et intérêts humains.

Dans son rapport intitulé «*Lignes directrices éthiques sur l'utilisation de l'intelligence artificielle (IA) et des données dans l'enseignement et l'apprentissage à l'intention des éducateurs*», la Commission européenne a défini plusieurs critères devant garantir que l'IA est utilisée de manière responsable et éthique dans les salles de classe.

(18) <https://www.rathenau.nl/nl/digitalisering/zo-staat-ai-op-de-agenda-van-de-eu>.

IV. ETISCHE RICHTSNOEREN VOOR ARTIFICIËLE INTELLIGENTIE (AI)

A. Op Europees niveau

De Europese Commissie heeft reeds verschillende initiatieven ondernomen om onderzoek naar en implementatie van AI op verschillende niveaus te bevorderen. Dit uit zich in de volgende initiatieven (18).

Om de krachten te bundelen en synergieën te creëren, heeft de Commissie samengewerkt met vijftientig lidstaten om tot een gezamenlijke verklaring te komen met betrekking tot AI.

Daarnaast heeft de Europese Commissie de *AI Alliance* opgericht, een platform dat als doel heeft om expertise, kennis en middelen te verzamelen en te delen, met als focus het bevorderen van AI-ontwikkeling en -implementatie binnen Europa.

De Commissie heeft ook een belangrijke mededeling gepubliceerd over artificiële intelligentie. Deze mededeling schetst de strategische visie van de Europese Unie over AI en benadrukt de cruciale rol ervan in de toekomst van Europa.

Om deze visie in daden om te zetten, heeft de Europese Commissie een gecoördineerd plan inzake artificiële intelligentie uitgewerkt. Dit plan legt de basis voor gezamenlijke inspanningen om de ontwikkeling, de implementatie en de regulering van AI-technologieën in Europa te bevorderen.

Als onderdeel van haar toewijding aan ethisch en verantwoord gebruik van AI, heeft de Europese Commissie ook een conceptgids voor ethische richtlijnen ontwikkeld. Deze gids bevat richtlijnen en principes die moeten worden nageleefd bij de ontwikkeling en de toepassing van AI, met het oog op het waarborgen van menselijke waarden en belangen.

De Europese Commissie heeft in het rapport «*Ethische richtsnoeren voor het gebruik van artificiële intelligentie en data bij onderwijs en leren voor onderwijsactoren*» verschillende criteria vastgelegd die ervoor moeten zorgen dat AI binnen de klaslokalen op een verantwoorde en ethische manier gebruikt wordt.

(18) <https://www.rathenau.nl/nl/digitalisering/zo-staat-ai-op-de-agenda-van-de-eu>.

Concrètement, ces lignes directrices s'articulent autour des sept axes suivants:

- *l'action humaine et le contrôle humain*, ce qui inclut les droits fondamentaux, les droits de l'enfant, le facteur humain et le contrôle humain;
- *la transparence*, ce qui inclut la traçabilité, l'explicabilité et la communication;
- *la diversité, la non-discrimination et l'équité*, ce qui inclut l'accessibilité, la conception universelle, l'absence de partialité et la participation des parties prenantes, qui permettent une utilisation indépendante de l'âge, du sexe, des capacités ou des caractéristiques, avec une attention particulière pour les apprenants ayant des besoins spécifiques;
- *le bien-être sociétal et environnemental*, ce qui inclut la durabilité et le respect de l'environnement, l'incidence sociale, la société et la démocratie;
- *le respect de la vie privée et la gouvernance des données*, ce qui inclut le respect de la vie privée, la qualité et l'intégrité des données, et l'accès aux données;
- *la robustesse technique et la sécurité*, ce qui inclut la résistance aux attaques, la sécurité et la sûreté générale, la précision, la fiabilité et la reproductibilité;
- *la responsabilisation*, ce qui inclut l'auditabilité, l'atténuation et la communication des incidences négatives, les arbitrages et les recours. Les considérations et les exigences peuvent aider les éducateurs (acteurs de l'enseignement), les chefs d'établissement et les fournisseurs de technologies à évaluer correctement l'incidence, à faire face aux dangers potentiels et à comprendre les avantages d'un système d'IA déployé et utilisé dans l'enseignement. En tant que telles, elles guident le développement, le déploiement et l'utilisation de systèmes d'IA dignes de confiance (19).

B. Au niveau national

AI 4 Belgium est une initiative fondée sur une approche communautaire. Elle a pour vocation de permettre aux personnes et aux organisations belges de saisir les opportunités de l'IA tout en facilitant de manière responsable la transition en cours. *AI 4 Belgium* a pour principale

Concreet komt het neer op de volgende zeven punten:

- *menselijke agency en toezicht*, met inbegrip van de grondrechten, de rechten van het kind, menselijke agency en menselijk toezicht;
- *transparantie*, met inbegrip van traceerbaarheid, verklaarbaarheid en communicatie;
- *diversiteit, non-discriminatie en rechtvaardigheid*, met inbegrip van toegankelijkheid, universeel ontwerp, het vermijden van onrechtvaardige vertekening, en participatie van belanghebbenden, waarmee gebruik mogelijk wordt gemaakt ongeacht leeftijd, geslacht, capaciteiten of kenmerken – met bijzondere aandacht voor leerlingen met speciale behoeften;
- *maatschappelijk en ecologisch welzijn*, met inbegrip van duurzaamheid en milieuvriendelijkheid, sociale impact, samenleving en democratie;
- *privacy en datagovernance*, met inbegrip van eerbiediging van de privacy, kwaliteit en integriteit van gegevens, en toegang tot gegevens;
- *technische robuustheid en veiligheid*, met inbegrip van weerbaarheid tegen aanvallen, beveiliging en algemene veiligheid, nauwkeurigheid, betrouwbaarheid en reproduceerbaarheid;
- *verantwoording*, met inbegrip van controleerbaarheid, minimalisering en verslaglegging van negatieve gevolgen, afwegingen en beroep. De overwegingen en vereisten kunnen onderwijsactoren, schoolleiders en technologieleveranciers helpen de impact adequaat te beoordelen, de potentiële risico's aan te pakken en de voordelen van een AI-systeem dat in het onderwijs wordt ingezet en gebruikt te benutten. Als zodanig sturen zij de ontwikkeling, de uitrol en het gebruik van betrouwbare AI-systemen (19).

B. Op nationaal niveau

AI 4 Belgium is een initiatief dat gestoeld is op een gemeenschapsgedreven aanpak. Het wezenlijke doel van dit initiatief bestaat erin om individuen en organisaties binnen België in staat te stellen de kansen die door AI worden geboden te benutten, terwijl de veranderingen

(19) https://learning-corner.learning.europa.eu/learning-materials/use-artificial-intelligence-ai-and-data-teaching-and-learning_fr.

(19) https://learning-corner.learning.europa.eu/learning-materials/use-artificial-intelligence-ai-and-data-teaching-and-learning_nl.

ambition de positionner la Belgique dans le paysage européen plus large de l'intelligence artificielle.

L'un des axes d'action d'*AI 4 Belgium* est de «renforcer le développement d'une infrastructure IA nationale et européenne compétitive, qui puisse activer l'IA chez tous les intervenants, y compris les services publics, les PME, les start-up, la recherche et la formation» (20).

C. Au niveau flamand et francophone

En Flandre et en Fédération Wallonie-Bruxelles, on suit les sept conditions clés définies par l'UE (21).

Les deux Communautés estiment qu'il est important de tenir compte de ces conditions si l'on envisage de mettre en œuvre un programme d'IA dans un établissement d'enseignement. Ces conditions donnent une orientation pour pouvoir exploiter tout le potentiel de l'IA tout en évitant les pièges éventuels. En respectant ces conditions, les établissements d'enseignement peuvent tirer parti des avantages de l'IA sans perdre de vue ses possibles implications négatives.

IV. QUE DEMANDONS-NOUS?

Les possibilités que l'IA recèle pour notre enseignement sont quasi illimitées et comparables aux nombreuses opportunités qu'elle a créées dans d'autres secteurs. Mais, outre ces perspectives prometteuses, l'IA comporte aussi des risques intrinsèques et des défis complexes. Il est donc capital d'intégrer cette technologie dans nos pratiques éducatives de manière réfléchie et stratégique.

(20) https://ai4be-5cbe3b103426b742c41e-endpoint.azureedge.net/blobai-4be0ee5d41015/wp-content/uploads/2023/01/Plan_national_de_convergence_pour_le_de%CC%81veloppement_de_lintelligence_artificielle.pdf.

(21) <https://www.vlaanderen.be/kenniscentrum-digisprong/themas/innovatie/artificiele-intelligentie/7-sleutelvoorwaarden-voor-betrouwbare-ai>. Voir le Service général du numérique éducatif (SGNE) qui s'occupe de mettre en œuvre le chantier «réussir la transition numérique à l'école». À titre d'exemple, les sept points relevés par l'UE sont repris dans le guide à destination des enseignants sur les incidences de *ChatGPT* sur l'enseignement.

die voortvloeien uit de toepassing van AI op verantwoorde wijze worden begeleid. De voornaamste ambitie van *AI 4 Belgium* ligt in het positioneren van België binnen het bredere Europese landschap van artificiële intelligentie.

Één van de actiepunten van *AI 4 Belgium* is «de ontwikkeling versterken van een competitieve nationale en Europese AI-infrastructuur, die AI bij alle betrokken partijen kan activeren, inclusief overheidsdiensten, kmo's, *start-ups*, onderzoek en onderwijs» (20).

C. Op Vlaams en Franstalig niveau

Op Vlaams niveau en in de Franse Gemeenschap volgt men dezelfde zeven sleutelvoorwaarden die de EU heeft vastgelegd (21).

Beide Gemeenschappen achten het belangrijk om bij de overweging om een AI-programma binnen de onderwijsinstelling te implementeren, deze kernvoorwaarden in gedachten te houden. Ze fungeren als een richtsnoer om het volledige potentieel van AI te benutten, terwijl potentiële valkuilen worden vermeden. Door aan deze voorwaarden te voldoen, kunnen scholen en onderwijsinstellingen de voordelen van AI daadwerkelijk verwezenlijken zonder de mogelijk negatieve implicaties uit het oog te verliezen.

V. WAT VRAGEN WIJ?

De kansen die AI voor ons onderwijs biedt, zijn haast onbegrensd en vergelijkbaar met de vele mogelijkheden die het in andere sectoren heeft ontketend. Echter, te midden van deze veelbelovende vooruitzichten, brengt AI ook inherente risico's en complexe uitdagingen met zich mee. Het is daarom van uiterst belang dat we deze technologie op een doordachte en strategische manier in onze onderwijspraktijk integreren.

(20) https://ai4belgium.be/wp-content/uploads/2023/01/Nationaal_convergentieplan_voor_de_ontwikkeling_van_artificiele_intelligentie.pdf.

(21) <https://www.vlaanderen.be/kenniscentrum-digisprong/themas/innovatie/artificiele-intelligentie/7-sleutelvoorwaarden-voor-betrouwbare-ai>. Zie de *Service général du numérique* (SGNE) die het project «réussir la transition numérique à l'école» uitvoert. De zeven door de EU aangestipte punten bijvoorbeeld staan in de gids voor de leerkrachten over de gevolgen van *ChatGPT* voor het onderwijs.

Il faut aussi garder à l'esprit que cette technologie peut faire l'objet d'une utilisation abusive et soulever, dans le futur, des questions en ce qui concerne le respect de la vie privée et la responsabilité. C'est une préoccupation qui a déjà émergé lorsqu'il s'est avéré que de grandes entreprises technologiques actives dans le secteur des médias sociaux avaient enfreint la loi à plusieurs reprises.

La possibilité que ce risque puisse s'étendre aujourd'hui à l'enseignement, pilier essentiel de notre démocratie, est donc inquiétante. Le fait d'abandonner un pan entier de nos processus éducatifs aux mains des grandes entreprises et de programmeurs, comme ce fut le cas pour les médias sociaux, est aussi préoccupant.

En outre, l'utilisation irréfléchie de l'IA peut causer des problèmes qui peuvent conduire rapidement à des pratiques inéquitables, déraisonnables, dangereuses, voire discriminatoires.

En réaction, nous proposons une politique holistique en six axes. Avant de mettre en œuvre un système d'IA, il faut définir des mesures stratégiques et des procédures à l'échelle de l'établissement d'enseignement. On disposera ainsi d'un cadre pour pouvoir cerner les attentes et élaborer des directives en vue de traiter les problèmes éventuels de manière cohérente. Ces mesures portent entre autres sur les éléments suivants:

- l'acquisition de programmes d'IA fiables et axés sur le facteur humain;
- la mise en œuvre d'un système de contrôle humain;
- la garantie que les données d'entrée sont pertinentes au regard de la destination du système d'IA;
- la dispensation d'une formation du personnel appropriée;
- la surveillance des performances du système d'IA et la mise en œuvre de mesures correctrices si nécessaire;
- le respect des obligations pertinentes du règlement général sur la protection des données (RGPD), y compris la réalisation d'une analyse d'impact de la protection des données.

Avec cette approche approfondie, nous pouvons tendre vers une mise en œuvre responsable et éthique de l'IA

Er bestaat echter ook een zorgwekkend vooruitzicht van potentieel misbruik, gepaard met vraagstukken rond privacy en verantwoordelijkheid in de toekomst. Dit is een kwestie die we eerder hebben gezien bij grote technologiebedrijven die herhaaldelijk wetten hebben overtreden, met name in de context van sociale media.

Het is dan ook zorgwekkend dat dit risico zich nu uitstrekt naar het onderwijs, een essentieel fundament van onze democratie. Het is verontrustend om een aanzienlijk deel van onze onderwijsprocessen in handen te geven van grote bedrijven en programmeurs, zoals eerder is gebeurd met sociale media.

Daarnaast kan ondoordacht gebruik van AI problemen opleveren die snel tot oneerlijke, onredelijke, gevaarlijke en zelfs discriminerende praktijken leiden.

Als reactie hierop introduceren we een holistisch zes-sporenbeleid. Voordat een AI-systeem wordt geïmplementeerd, dienen er schoolbrede beleidsmaatregelen en procedures te worden vastgesteld. Dit biedt een kader om verwachtingen te formuleren en richtlijnen te creëren voor een consistente aanpak van eventuele problemen. Deze maatregelen hebben onder andere betrekking op:

- het aanschaffen van betrouwbare en op de mens gerichte AI;
- het implementeren van menselijk toezicht;
- het waarborgen van relevante inputgegevens die passen bij het beoogde doel van het AI-systeem;
- het voorzien van geschikte training voor het personeel;
- het monitoren van de prestaties van het AI-systeem en het nemen van corrigerende maatregelen indien nodig;
- de naleving van de vereisten van de algemene verordening gegevensbescherming (AVG), inclusief het uitvoeren van een beoordeling van gegevensbeschermingseffecten.

Met deze grondige aanpak kunnen we naar een verantwoorde en ethische implementatie van AI in het

dans l'enseignement, tout en garantissant la protection de la vie privée des individus et l'intégrité du processus d'apprentissage.

L'enseignement est une compétence des Communautés. La sécurité, la protection de la vie privée et la technologie sont des compétences partagées entre les entités fédérées et le niveau fédéral. La présente proposition de résolution a donc bien un caractère transversal.

*
* *

onderwijs streven, waarbij de privacy van individuen en de integriteit van het leerproces worden gewaarborgd.

Onderwijs is een bevoegdheid van de Gemeenschappen. Veiligheid, privacy en technologie zijn gedeelde bevoegdheden tussen de deelstaten en het federale niveau. Dit voorstel van resolutie heeft aldus een transversaal karakter.

*
* *

PROPOSITION DE RÉOLUTION

Le Sénat,

A. considérant que le recours à l'intelligence artificielle (IA) en classe suscite un enthousiasme de plus en plus grand dans les milieux scientifique et pédagogique;

B. considérant que l'utilisation de l'IA dans l'enseignement n'est pas dénuée de risques importants, tant pour l'étude proprement dite que pour l'évaluation;

C. considérant qu'il existe un risque inhérent à la combinaison de l'IA et d'un régime autoritaire et qu'il s'impose dès lors d'élaborer d'urgence des lignes directrices éthiques;

D. considérant que l'IA repose sur des instructions, à savoir des algorithmes qui permettent aux ordinateurs d'exécuter diverses tâches;

E. considérant que l'apprentissage automatique (*machine learning*) utilise des données et des algorithmes afin de réaliser des tâches traditionnellement exécutées par l'homme, sans que des instructions strictes soient communiquées;

F. considérant que, dans un contexte plus complexe, l'algorithme de l'IA est parfois comparé à un réseau neuronal qui présente des similitudes avec les schémas de pensée humains;

G. considérant que les systèmes d'IA puisent des données dans différentes sources et peuvent contribuer à améliorer l'enseignement en facilitant la compréhension des besoins d'apprentissage, les expériences d'apprentissage personnalisées et le processus décisionnel;

H. considérant que les systèmes d'IA peuvent faciliter l'enseignement et l'apprentissage;

I. considérant que «l'IA éthique» vise la mise au point, le déploiement et l'utilisation d'une intelligence artificielle qui respecte des normes éthiques et des valeurs fondamentales;

J. considérant que l'utilisation de systèmes d'IA offre des avantages pour les apprenants, car elle encourage un apprentissage personnalisé dans le cadre duquel les besoins d'apprentissage individuels sont mieux compris et adaptés;

VOORSTEL VAN RESOLUTIE

De Senaat,

A. overwegende dat binnen de wetenschap en de pedagogie men in groeiende mate enthousiast is omtrent het gebruik van artificiële intelligentie (AI) in de klaslokalen;

B. overwegende dat het gebruik van AI in het onderwijs aanzienlijke risico's met zich kan meebrengen, zowel tijdens het studeren als bij het beoordelen;

C. overwegende dat er een gevaar schuilt in de combinatie van AI en autoritaire regimes, waarvoor ethische richtlijnen dringend noodzakelijk zijn;

D. overwegende dat AI draait om geïnstrueerde handelingen, namelijk algoritmes die computers helpen bij het uitvoeren van diverse taken;

E. overwegende dat machinaal leren (*machine learning*) gegevens en algoritmes gebruikt om taken uit te voeren die normaal door mensen uitgevoerd worden, zonder rigide instructies;

F. overwegende dat in een complexere context het AI-algoritme soms wordt vergeleken met een neurale netwerk dat op menselijke denkpatronen lijkt;

G. overwegende dat AI-systemen gegevens uit verschillende bronnen halen en het onderwijs kunnen helpen verbeteren door leerbehoeften, persoonlijke leerervaringen en betere besluitvorming te implementeren;

H. overwegende dat AI-systemen lesgeven en leren kunnen vergemakkelijken;

I. overwegende dat «Ethische AI» zorgt voor ontwikkeling, uitrol en gebruik van AI in lijn met ethische normen en kernwaarden;

J. overwegende dat het gebruik van AI-systemen voor leerlingen voordelen biedt door gepersonaliseerd leren te benadrukken, waarbij individuele leerbehoeften beter begrepen en aangepast worden;

K. considérant que les apprenants peuvent apprendre à leur propre rythme en fonction de leurs points forts et points faibles, en bénéficiant d'une évaluation automatique qui leur fournit un *feed-back* immédiat et leur permet de mieux assimiler la matière enseignée;

L. considérant que l'IA représente un gain de temps pour les enseignants grâce à l'automatisation des tâches administratives, ce qui leur permet de consacrer davantage de temps à l'interaction avec les apprenants;

M. considérant qu'une étude de la société de consultance *McKinsey* montre que les enseignants consacrent 20 à 40 % de leur temps de travail à des tâches qui pourraient être automatisées;

N. considérant qu'on évoque souvent le problème de la «boîte noire», le fonctionnement de l'IA n'étant pas bien connu;

O. considérant que les performances de l'IA dépendent de la qualité des données qui l'alimentent (degré d'entraînement) et qu'il est important de mettre en balance, d'une part, l'introduction de données et, d'autre part, la vie privée et la sécurité;

P. considérant que diverses initiatives ont déjà été prises en matière d'IA au niveau européen;

Q. vu les lignes directrices éthiques sur l'utilisation de l'intelligence artificielle et des données dans l'enseignement et l'apprentissage à l'intention des éducateurs, édictées par la Commission européenne en 2022;

R. considérant que les Régions et Communautés du pays sont amenées à suivre ces lignes directrices éthiques;

S. considérant qu'au niveau national, l'initiative *AI 4 Belgium*, entre autres, promeut l'implémentation de l'IA dans différents secteurs, via le Plan national de convergence du 28 octobre 2022 pour le développement de l'intelligence artificielle, qui vise à promouvoir un modèle éthique, équilibré entre l'innovation et la protection des droits fondamentaux;

T. vu le rapport d'information relatif à la nécessaire collaboration entre l'État fédéral et les entités fédérées en ce qui concerne les retombées, les opportunités, les potentialités et les risques de la «société intelligente» numérique, et plus particulièrement les chapitres relatifs à l'éthique et aux droits de l'homme ainsi qu'à l'enseignement, approuvé par le Sénat le 29 mars 2019 (doc. Sénat, n° 6-413/3),

K. overwegende dat leerlingen op hun eigen tempo kunnen leren op basis van hun sterke en zwakke punten, met automatische beoordeling voor directe *feedback* en dieper begrip van de leerstof;

L. overwegende dat AI een tijdsbesparing betekent voor leerkrachten door automatisering van administratieve taken, waardoor ze meer tijd hebben voor interactie met leerlingen;

M. overwegende dat *McKinsey's* onderzoek aantoont dat leerkrachten 20 tot 40 % van hun werktijd besteden aan taken die geautomatiseerd kunnen worden;

N. overwegende dat er vaak sprake is van een «*black box*» probleem, waarbij men niet goed weet hoe AI werkt;

O. overwegende dat AI slechts zo goed is als de data waarop het getraind is en de afweging tussen de invoer van data enerzijds en privacy en veiligheid anderzijds belangrijk is;

P. overwegende dat men op Europees niveau al verschillende initiatieven rond AI genomen heeft;

Q. gelet op de ethische richtsnoeren die de Europese Commissie in 2022 heeft uitgevaardigd voor het gebruik van artificiële intelligentie en data bij onderwijzen en leren voor onderwijsactoren;

R. overwegende dat de Gewesten en Gemeenschappen van het land deze ethische richtsnoeren moeten volgen;

S. overwegende dat op nationaal niveau onder andere *AI 4 Belgium* de implementatie van AI in diverse sectoren promoot, via het nationaal convergentieplan van 28 oktober 2022 voor de ontwikkeling van artificiële intelligentie, dat ertoe strekt een ethisch model te bevorderen dat een evenwicht biedt tussen innovatie en bescherming van de fundamentele rechten;

T. gelet op het informatieverslag betreffende de noodzakelijke samenwerking tussen de Federale Staat en de deelstaten inzake de impact, de kansen en mogelijkheden en de risico's van de digitale «slimme samenleving», en meer bepaald op de hoofdstukken over ethiek en mensenrechten en over het onderwijs, door de Senaat aangenomen op 29 maart 2019 (doc. Senaat, nr. 6-413/3),

Demande à tous les gouvernements compétents dans ce domaine:

- 1) de poursuivre la mise en œuvre des mesures actuelles visant à promouvoir la recherche et le développement de l'IA, comme le prévoient les notes de politique générale et accords de gouvernement;
- 2) de lancer une étude sur les moyens d'encourager les entreprises (publiques) et les établissements d'enseignement qui envisagent d'acquérir une technologie d'IA à rechercher un label ou une certification confirmant que le système d'IA en question répond à des normes éthiques et axées sur l'humain. Cette certification devrait être délivrée par une instance indépendante et indiquerait que le système d'IA convient à une utilisation dans des contextes éducatifs;
- 3) de mettre en place une autorité indépendante de contrôle des algorithmes qui sera chargée de veiller à ce que l'utilisation de l'IA au sein d'organismes publics respecte des directives éthiques, y compris, si possible, dans le domaine de l'enseignement. Il s'agira en l'occurrence de vérifier si les systèmes d'IA respectent les normes éthiques, les exigences de transparence et les règles en matière de respect de la vie privée;
- 4) de proposer, en vue d'une mise en œuvre responsable de l'IA, que les établissements d'enseignement et de formation, les entreprises et les institutions publiques aient l'obligation d'informer clairement les étudiants, les parents et les travailleurs sur le recours à l'IA dans les processus d'évaluation et d'apprentissage. Il s'agit en l'occurrence de préciser de quelle manière l'IA est appliquée, quelles données sont utilisées et de quelle manière l'IA intervient dans la prise de décisions;
- 5) d'encourager les établissements d'enseignement et de formation, les entreprises et les institutions publiques à procéder à des évaluations régulières de l'usage qu'ils font de l'IA. Ces évaluations devront porter sur l'efficacité, la sécurité et les aspects éthiques des technologies d'IA mises en œuvre et viseront à identifier les éventuelles adaptations nécessaires;
- 6) d'encourager les établissements d'enseignement à collaborer avec des experts de l'IA et des spécialistes de la protection de la vie privée afin de veiller à ce que la mise en œuvre de l'IA dans les classes se fasse conformément aux bonnes pratiques en matière de sécurité et de respect de la vie privée. Une formation continue des enseignants et des établissements est donc indispensable;

Vraagt alle hiertoe bevoegde regeringen om:

- 1) de huidige maatregelen die de ontwikkeling en het onderzoek van AI bevorderen, zoals vastgelegd in de beleids- en regeringsnota's, verder na te streven;
- 2) een onderzoek te initiëren naar het stimuleren van (openbare) bedrijven en onderwijsinstellingen die overwegen om AI-technologie aan te schaffen, hierbij te streven naar een label of certificering die bevestigt dat het betreffende AI-systeem aan ethische en mensgerichte normen voldoet. Deze certificering zou moeten worden afgegeven door een onafhankelijke instantie en zou aangeven dat het AI-systeem geschikt is voor gebruik in educatieve contexten;
- 3) een onafhankelijk algoritmetoezichthouder aan te stellen die erop toeziet dat het gebruik van AI bij openbare instanties volgens ethische richtlijnen verloopt, ook waar mogelijk op het gebied van onderwijs. Hierbij zal worden gecontroleerd of de AI-systemen in lijn zijn met ethische normen, transparantievereisten en privacyvoorschriften;
- 4) met het oog op een verantwoordelijke inzet van AI, voor te stellen dat onderwijs- en opleidingsinstellingen, bedrijven en openbare instellingen een meldingsplicht hebben om duidelijke informatie te verschaffen aan studenten, ouders en werknemers over het gebruik van AI bij beoordelingsprocessen en leerprocessen. Hierbij wordt ingegaan op de manier waarop AI wordt toegepast, welke gegevens worden gebruikt en hoe beslissingen worden genomen met behulp van AI;
- 5) onderwijs- en opleidingsinstellingen, bedrijven en openbare instellingen aan te moedigen om regelmatige beoordelingen uit te voeren bij het gebruik van AI. Deze beoordelingen dienen te focussen op de effectiviteit, veiligheid en ethische aspecten van de ingezette AI-technologieën, en om eventuele noodzakelijke aanpassingen te identificeren;
- 6) onderwijsinstellingen te stimuleren om samen te werken met AI-experts en privacydeskundigen om ervoor te zorgen dat de implementatie van AI in de klaslokalen in lijn is met de *best practices* op het gebied van veiligheid en privacy. Een constante bijscholing van docenten en instellingen is dan ook onontbeerlijk;

7) de garantir en permanence le respect de la vie privée et la protection des données des étudiants, des parents et des travailleurs, y compris en ce qui concerne l'IA. Cela permettra de faire en sorte que le recours à l'IA dans l'enseignement se fasse dans le respect du règlement général sur la protection des données (RGPD) et protège les intérêts de toutes les parties concernées.

Le 4 octobre 2023

7) blijvend de privacy en gegevensbescherming van studenten, ouders en werknemers te waarborgen, ook op het gebied van AI. Dit zorgt ervoor dat de inzet van AI in het onderwijs in overeenstemming is met de algemene verordening gegevensbescherming (AVG) en het de belangen van alle betrokkenen beschermt.

4 oktober 2023

Rik DAEMS.

Fatima AHALLOUCH.

Gaëtan VAN GOIDSENHOVEN.