

Dans les coulisses de l'«intelligence artificielle»: un regard sociologique sur les dangers et les potentiels des logiciels d'apprentissage adaptatifs

KENNETH HORVATH, ANDREA ISABEL FREI & MARIO STEINBERG

L'emploi de l'intelligence artificielle (IA) dans des contextes éducatifs soulève des questions fondamentales par rapport à l'équité. Ainsi, on associe souvent à l'IA la promesse de pouvoir agencer des processus de formation plus équitables, car radicalement plus personnalisés. Face à cela, il y a un risque de discrimination algorithmique (la reproduction d'inégalités existantes par les technologies éducatives). Dans ce contexte, l'article présente un projet de recherche en cours qui examine l'interaction entre les champs technologique et pédagogique dans le développement et la mise en œuvre de technologies de formation assistées par l'IA. Les premiers constats de ce projet soulignent la nécessité de s'interroger sur le monitoring empirique de l'IA et d'une réflexion collaborative sur celle-ci dans la formation.

Au plus tard depuis le battage médiatique autour de la production automatique de textes par le bot «Chat GPT» au début de cette année, le débat sur le rôle futur de l'intelligence artificielle dans le contexte éducatif s'est échauffé. L'impressionnante capacité à imiter les genres et à associer des contenus de manière plausible a focalisé ces discussions sur le problème de «l'authenticité et de la contrefaçon» et sur l'utilisation pédagogique appropriée de ces technologies. On a ici généralement perdu de vue un autre aspect: les questions fondamentales relatives à l'équité qui vont de pair avec l'emploi de l'IA dans le contexte éducatif.

FSEA (Éd.): Education Permanente 2023-1, Revue suisse pour la formation continue, www.ep-web.ch/f



Ces questions découlent de la promesse centrale associée à l'IA dans la formation: permettre davantage d'équité dans la formation grâce à une personnalisation plus radicale des processus d'apprentissage. Face à ces espoirs, il y a de sérieuses mises en garde contre les dangers d'une discrimination algorithmique: l'IA identifie des schémas dans de grands ensembles de données historiques et est donc susceptible de reproduire des mécanismes de préjugés existants.

Ce problème constitue le point de départ d'un projet de recherche financé par le Fonds national suisse (FNS), qui a démarré récemment à la Haute école pédagogique de Zurich. Dans cet article, nous énonçons les questions essentielles sur lesquelles porte le projet, et donnons un aperçu des constats et apports qui pourraient en résulter. Il se concentre sur les classifications riches en exigences et en conséquence qui vont de pair avec l'intelligence artificielle et l'apprentissage automatique (*machine learning*). Afin de bien comprendre ces classifications et leurs implications, il faut tout d'abord décrypter l'interaction des champs pédagogique et technologique dans le développement et l'utilisation de ces technologies de formation.

Nous commençons par donner ci-après un aperçu d'ensemble des formes et des fonctions actuelles des technologies de formation assistées par l'IA. Sur cette base, nous esquissons ensuite les hypothèses directrices, les enjeux et les résultats attendus du projet «Algorithmic Sorting in Education»¹.

Qu'est-ce exactement que l'«intelligence artificielle» et que fait-elle?

Dans les débats médiatiques, l'IA est souvent présentée comme la vision d'avenir ultime de notre ère numérique (Jasanoff, 2015; Selwyn, 2022). Cette manière de circonscrire les choses détourne l'attention du fait qu'il est déjà question d'intelligence artificielle depuis de longues décennies – et que les visions fondamentales, précisément par rapport à son utilisation dans les domaines éducatifs, ont étonnamment peu changé. La vision centrale est toujours celle d'une personnalisation automatisée et radicale de l'apprentissage (Skinner, 1958; du Boulay, 2019).

¹ <https://data.snf.ch/grants/grant/204737>

Les variantes actuelles de l'intelligence artificielle se distinguent avant tout par la manière dont elles sont censées concrétiser cette vision. Essentiellement, il s'agit de déceler des schémas (régularités statistiques sans aucune conjecture de lien de cause à effet) dans de grands volumes de données en vue de catégoriser de nouvelles données. Avec les immenses quantités de données de sources les plus diverses disponibles aujourd'hui, la classification assistée par l'IA est utilisée à des fins multiples. Ainsi, on déduira des pronostics de succès individuels à partir de données sur le comportement d'apprentissage et d'examen, cela afin d'identifier de façon précoce des parcours d'apprentissage supposément problématiques. Mais l'IA peut également être utilisée, par exemple, pour l'évaluation automatique de textes. À l'arrière-plan, il y a ici toujours une comparaison de nouveaux cas avec des cas plus anciens (consignés dans les données existantes). La notion ambitieuse d'*intelligence* est donc réduite à une seule fonction cognitive: la classification (Horvath, 2020).

Ce que l'on appelle aujourd'hui l'IA se distingue d'autres algorithmes par la manière implicite et indirecte de parvenir à de telles classifications (Dourish, 2016). Au départ, un algorithme n'est rien d'autre qu'un ensemble de règles pour l'accomplissement de certaines tâches. Les algorithmes «traditionnels» se composent par exemple de règles de la forme «si... alors...», dans lesquelles les conditions et instructions pour l'action sont explicitement indiquées (Neyland & Möllers, 2017).

Face à cela, les algorithmes assistés par l'IA tels qu'on les entend aujourd'hui dans le langage courant s'alignent sur l'idée de l'apprentissage automatique (*machine learning*): l'algorithme est «entraîné» au moyen d'un vaste ensemble de données. Suivant des instructions d'action déterminées, il y cherche des schémas saillants qui apparaissent en lien avec un résultat donné. Les algorithmes IA «apprennent» donc tout d'abord, au moyen d'exemples qui leur sont présentés, comment ceux-ci sont à évaluer. Les schémas décelés sont appliqués comme règles de décision aux nouvelles données.

Même les programmeurs et programmatrices de tels algorithmes ne savent généralement pas vraiment quels facteurs sont déterminants à l'arrière-plan, dans quelle combinaison et avec quelle pondération. On sait seulement quelles données sont à la disposition de l'algorithme, mais pas comment elles sont exactement utilisées pour la classification des cas. Les algorithmes assistés par l'IA ressemblent ainsi au travail de catégorisation par des expert·e·s qui, sur la base de leur longue expérience, décèlent plus ou moins inconsciemment des schémas et peuvent y réagir. Ils «sentent» à quel type de cas ils ont affaire.

Ce n'est pas seulement le terme d'intelligence artificielle qui est courant depuis longtemps déjà – les techniques concrètes qui sont employées

aujourd'hui, elles aussi, sont loin d'être nouvelles. Qu'il s'agisse de réseaux neuronaux ou d'arbres de décision – dans leurs grandes lignes, les procédés utilisés par l'IA moderne sont établis depuis des dizaines d'années déjà. Mais ce qui est nouveau, c'est la quantité des données au sein desquelles les algorithmes peuvent aujourd'hui chercher des schémas et la rapidité avec laquelle ils le font grâce aux performances actuelles des ordinateurs.

Domaines d'utilisation de l'IA dans la formation (continue)

Le potentiel de l'IA réside dans la multiplicité des tâches pour lesquelles elle peut être utilisée. Notamment dans la formation continue, de telles technologies sont déjà employées régulièrement dans des endroits divers (Bartolomé et al., 2018; Gibson, 2017; Kirkwood & Price, 2014; Prinsloo, 2020). Parmi les applications existantes comptent:

- Des outils pour pronostiquer les succès d'apprentissage (Riazy et al., 2020): dans l'espace anglophone, l'IA est déjà utilisée dans de nombreuses hautes écoles pour des systèmes d'alerte précoce. Ces programmes utilisent de multiples données pour une classification des étudiant·e·s relative à leur comportement d'études ultérieur et à leur réussite attendue (en plus des résultats d'examen, aussi le nombre d'accès aux contenus, les durées de traitement, mais également des données de suivi oculaire (*Eye-Tracking*), etc.).
- Des systèmes tutoriels «intelligents» (Ma et al., 2014): un autre domaine d'application réside dans les logiciels qui agencent les contenus et voies d'apprentissage de façon adaptative. L'IA est ici utilisée d'une part pour une classification des utilisateurs·trices en fonction de leurs forces, faiblesses, besoins et intérêts, et d'autre part pour ordonnancer les contenus et les tâches en fonction de leur «résolvabilité» pour différents groupes d'apprenant·e·s.
- Des tableaux de bord «intelligents» (Jarke & Macgilchrist, 2021; Verbert et al., 2020): l'IA est par ailleurs utilisée dans des tableaux de bord au sein desquels les apprenant·e·s obtiennent des informations en temps réel à propos de leurs groupes d'apprentissage. Ici également, l'IA sert par exemple à identifier des situations d'apprentissage «critiques» et des étudiant·e·s en risque d'échec.
- Des systèmes d'aide à la rédaction (Dixon-Román et al., 2020): un autre domaine d'application de l'IA concerne l'évaluation automatisée de textes. La vision directrice est celle de programmes qui fournissent un feed-back formatif sur les textes – en partie déjà pendant le processus de rédaction.

La promesse: l'IA apporte davantage d'équité dans la formation

Divers arguments sont avancés en faveur de l'emploi de l'IA. À côté de l'augmentation de l'efficacité et de la réduction des coûts, les promesses d'équité jouent ici un rôle majeur (Dencik et al., 2019; Francis et al., 2020; Riazzy et al., 2020). L'IA devrait permettre de concevoir la formation de manière plus équitable, premièrement parce qu'elle est, contrairement à l'être humain, en mesure de porter un regard sans préjugés, «objectif», sur les apprenant·e·s, et deuxièmement précisément parce qu'elle permet l'adaptativité. Les offres de formation personnalisées sont par conséquent plus équitables, parce qu'elles permettent de tenir compte des besoins, des intérêts et des talents de tout individu. L'IA devrait donc permettre une formation personnalisée de qualité «pour toutes et tous» et répondre ainsi en même temps aux besoins de main-d'œuvre flexible, autonome et compétente d'un (futur) marché du travail.

Les récits de ce type sont largement répandus (Horvath et al., 2023). Mais cet espoir est un peu prématuré à plusieurs égards. Une préoccupation concerne l'assimilation sans réserve de l'adaptativité et de l'équité de formation (Selwyn, 2022). Or, la formation personnalisée n'est pas nécessairement plus équitable. La personnalisation des voies d'apprentissage peut tout au contraire renforcer des schémas de sélectivité sociale de la formation. Ainsi, des boucles de rétroaction lourdes de conséquences peuvent par exemple s'établir entre l'autoperception de l'apprenant·e et les classifications algorithmiques. Si l'on a moins confiance en sa propre performance, on traitera les tâches plus lentement; cette information est évaluée par l'algorithme comme une faiblesse de performance; cette classification a pour conséquence que des tâches plus complexes ne seront pas attribuées à la personne en question, et ainsi de suite.

Deuxièmement, il faut se demander ce qu'une IA personnalise exactement et sur quelle base (Bartolomé et al., 2018). On parle souvent d'une personnalisation s'alignant sur les besoins, les intérêts et les talents. Dans la pratique, on remarque toutefois que l'on ne récolte presque jamais d'informations sur ces aspects (Horvath et al., 2023). Or, un algorithme ne peut traiter pour la détection de schémas que ce dont il dispose sous forme de données. Les programmes assistés par l'IA d'usage courant à l'heure actuelle se limitent pour la plupart à des données de performance, telles que la justesse des réponses ou le temps de traitement des tâches.

Troisièmement, la promesse d'une liberté de préjugés doit également être considérée avec prudence. Car l'IA peut très bien discriminer. Afin que l'on puisse comprendre les défis posés par l'IA aux institutions de formation et à la pratique pédagogique, ce phénomène revêt une importance majeure.

Le danger: l'IA entraîne une discrimination algorithmique

Reconnaître que les algorithmes d'intelligence artificielle puissent avoir un effet «discriminatoire» n'est pas un fait nouveau (Beer, 2017; Eubanks, 2017; O'Neil 2016; Noble, 2018). Il y a d'innombrables exemples de distorsions concernant le genre ou la couleur de peau – algorithmes censés évaluer pour la justice aux États-Unis le risque de récidive des personnes condamnées et qui défavorisent systématiquement les personnes non blanches, schémas sexistes d'algorithmes de recherche assistés par l'IA ou «correction» sociale-ment biaisée de résultats d'examens au Royaume-Uni lorsque l'on a craint au début de la pandémie de Covid-19 une évaluation trop indulgente des examens de fin d'études des élèves (Williamson, 2021; Selwyn, 2022).

De telles formes de «discrimination algorithmique» ont une raison toute simple: les algorithmes assistés par l'IA reproduisent les schémas qu'ils trouvent dans les données existantes. Si l'on a donc, par le passé, systématiquement utilisé la couleur de la peau comme indicateur de l'aptitude à la réhabilitation, de la solvabilité, etc., un algorithme qui n'est pas contrôlé de manière appropriée et n'intervient pas de manière correctrice va forcément perpétuer ces schémas. L'un des arguments les plus forts en faveur de l'emploi de l'IA devient ainsi un argument contre celui-ci.

La gestion de ce mécanisme pose aux développeurs-euses un problème de conception technique de la programmation: comment doit-on et peut-on, dans le travail quotidien avec et sur l'IA, gérer le risque des discriminations algorithmiques? L'une des marges de manœuvre essentielles pour la décision et la conception relève de la question de savoir comment les algorithmes doivent traiter les variables de contexte social (Whitman, 2020).

L'aspect décisif est que les développeurs-euses doivent prendre des décisions potentiellement lourdes de conséquences, mais dont les impacts précis sont difficiles à évaluer. Ainsi se posent deux questions: 1. Sur quelle base ces décisions sont-elles prises – quels aspects sont-ils pris en compte, quelles sont les raisons déterminantes? 2. Quels effets les technologies une fois conçues déploient-elles au fil de leur mise en œuvre pratique dans des contextes de formation concrets? C'est sur ces questions que se penche le projet «Algorithmic Sorting in Education».

Le problème étudié: présupposés implicites et incontrôlés

Pour comprendre les défis et les conséquences de l'IA dans des contextes éducatifs, il est central de savoir que l'IA prend des décisions en permanence – et ce sur la base de règles implicites et inconnues (Al-Amoudi & Latsis, 2019). Le contrôle de ces décisions algorithmiques au niveau pédagogique, technologique et de la politique éducative doit donc s'attacher à définir et à ap-

profondir les conditions cadres pour le système de décision qu'est l'algorithme. Quelles données doivent être utilisées? Quelle méthode ou quel type d'algorithme veut-on utiliser exactement? Quelles mesures de sécurité et de monitoring faut-il prendre? Selon quels critères le travail de l'IA doit-il être apprécié, et qui en décide?

Le projet de recherche mentionné étudie comment ces conditions cadres du développement et de l'utilisation de l'IA sont agencées dans des contextes de formation par les parties prenantes dans les champs technologique et pédagogique. À cet effet, il combine deux modules empiriques: 1. Dans des groupes de réflexion, on examine sur quels présupposés pédagogiques et technologiques s'appuient les acteurs et actrices dans la réflexion et les débats sur les triages algorithmiques dans la formation. 2. Dans des études de cas, on examine comment, sur la base de ces présomptions, on prend dans le quotidien professionnel des décisions concrètes sur l'agencement technologique et/ou l'utilisation pédagogique de technologies de formation assistées par l'IA. Ici, l'accent porte dans un premier temps sur l'école obligatoire, mais d'autres domaines éducatifs seront également pris en compte au fur et à mesure du projet, y compris la formation pour adultes et la formation continue.

Nos premières approches empiriques suggèrent déjà que l'on a souvent affaire – un peu comme pour l'IA elle-même – à des processus de décision et de conception plutôt cachés, sur la base de présomptions ad hoc implicites qui n'ont pas fait l'objet d'une réflexion consciente. Afin de rendre plus tangible l'importance de tels présupposés: de toute évidence, pour les programmeurs·trices, le développement de logiciels assistés par l'IA va de pair avec beaucoup de décisions technologiques au sens large – de la sélection d'un langage de programmation approprié à la conception d'une interface utilisateurs fonctionnelle. En outre, il est toutefois également indispensable de définir des aspects pédagogiques (Horvath et al., 2023):

- Une technologie éducative doit impérativement «mettre en scène» une théorie quelconque de ce qu'est l'apprentissage et de son fonctionnement. On ne doit donc s'attendre que dans de très rares cas à des théories élaborées et détaillées de l'apprentissage, mais une conception quelconque de l'apprentissage est inévitablement intégrée dans le code.
- Cela s'associe également à des décisions de valeur implicites. Il s'agit essentiellement de savoir en quoi consiste un «bon» apprentissage.
- Cette définition de la qualité de la formation est étroitement liée à des questions de pertinence: toute sélection de tâches et de contenus implique une détermination de ce qui peut et doit être considéré comme un contenu d'apprentissage «important».
- En parallèle, on sent ainsi toujours transparaître une certaine idée – réfléchie ou non – de quand l'éducation est «équitable» et quand les pro-

cessus d'apprentissage doivent être considérés comme inévitables. Dans la pratique, cette question devient souvent pertinente *ex negativo*: la technologie peut être jugée sans danger et être commercialisée tant qu'aucune réserve relative à l'équité n'est formulée.

- Enfin, toute technologie éducative implique une hypothèse sur la question de savoir où l'apprentissage a effectivement lieu au quotidien, et qui y contribue dans quel rôle.

Ces présupposés ne sont que très rarement négociés et déterminés de manière consciente. Ils viennent plutôt s'insinuer par des voies subtiles et indirectes dans le développement de la technologie. Ainsi, de nombreuses technologies éducatives assistées par l'IA reprennent des échelles et des items d'instruments de mesure psychométriques existants, tels qu'on les utilise par exemple dans les évaluations à grande échelle (Perrotta & Selwyn, 2020). Or, avec les échelles, on reprend forcément aussi les conceptions correspondantes du succès de formation et d'apprentissage.

Dans le champ pédagogique également, il faut compter sur des présupposés qui n'ont pas fait l'objet d'une réflexion suffisante – même si, dans ce cas, ils prennent plutôt la forme de représentations de ce qu'est réellement l'IA et de ce que font effectivement certaines technologies – ce qui a des conséquences plus importantes. Dans ce cas, le danger réside dans le fait que l'on estime sa propre compétence insuffisante pour une réflexion critique, et que l'on reprend les jugements de l'IA sans les remettre en question. Il est également possible que des enseignant·e·s refusent d'utiliser l'application – au titre de leur liberté professionnelle – ou que les technologies soient employées dans la pratique pédagogique autrement que prévu à l'origine.

Le challenge: comprendre l'interaction de la pédagogie et de la technologie

Notre projet de recherche est guidé par l'hypothèse qu'une compréhension approfondie de ces processus d'agencement et d'utilisation au quotidien, avec les valeurs et les hypothèses sous-jacentes, constitue une exigence de base pour une utilisation de l'IA sensible à la diversité et aux inégalités dans la formation.

Le projet vise à jeter un regard dans les coulisses des technologies éducatives assistées par l'IA – ou plus précisément: à déceler les processus d'évaluation, de réflexion, de négociation et de conception qui caractérisent le développement et l'utilisation des technologies assistées par l'IA, lorsque les acteurs·trices interviennent dans des situations très concrètes de leur quotidien professionnel dans les champs pédagogique et technologique. Nous supposons qu'il y a entre ces deux champs des rapports et des interactions multiples, mais qui ne sont guère compris jusqu'à présent.

Nos premières approches empiriques font d'ores et déjà ressortir qu'il faut s'attendre ici à des tensions lourdes de conséquences – et qu'il y a donc matière à réflexion et à discussion. Citons un exemple:

Pour le développement des technologies éducatives novatrices, les start up jouent un rôle de leader. Ces jeunes entreprises sans cesse confrontées au risque d'un échec sont soumises à une pression tout à fait spécifique d'agir. À première vue, les conceptions d'une «formation bonne et équitable» ne sont pas au premier plan; il s'agit avant tout de convaincre des individus disposant de moyens financiers importants («business angels») ou des fonds de placement de la viabilité commerciale d'une idée. Néanmoins, des présomptions relevant de la pédagogie et de la politique éducative font immanquablement partie de l'affaire: elles s'y intègrent sous forme d'hypothèses quant aux client·e·s potentiel·le·s et à leurs besoins et préférences, sous forme de suppositions quant aux stratégies de marketing utiles ou encore sous forme d'échelles et de tâches.

Déjà lors d'une première approche empirique, le contact riche en conflits de conceptions divergentes mais simultanées sur la qualité et l'équité de la formation saute aux yeux. Des récits qui s'articulent autour de la créativité, de l'individualité et de l'authenticité comme indice d'une formation «bonne et équitable» font face à des technologies qui visent avant tout la simplification, la standardisation et l'augmentation de l'efficacité.

Or, tant les environnements d'apprentissage autorégulés (qui s'alignent sur des récits axés sur la créativité) que les contextes d'apprentissage standardisés (qui suivent des logiques plus «industrialisées» d'une éducation bonne et équitable) peuvent générer une sélectivité sociale – mais par des voies très différentes: dans les contextes d'apprentissage standardisés, les défavorisations systématiques résultent de distorsions subtiles, mais systématiques, de formats d'examen standardisés; dans les environnements d'apprentissage autorégulés, elles découlent plutôt de formes d'autoclassification.

La recherche en sociologie de l'éducation des dernières dizaines d'années a fait ressortir ces différentes voies de préjudice systématique, qui peuvent donc faire l'objet d'une attention et d'une réflexion professionnelles. Les choses se compliquent lorsqu'une technologie inscrite dans un récit comme «axée sur la créativité» et «personnalisante» est effectivement utilisée dans des environnements (autorégulés) agencés en conséquence, mais que la technologie en soi classe et aborde les apprenant·e·s selon des logiques tout à fait différentes, mesurant de manière standardisée. Dans un tel cas de logiques «mélangées» et incontrôlées, on court le risque que des préjudices non intentionnels se produisent – qui sont toutefois nettement plus difficiles à détecter et à remettre en question que dans des environnements d'apprentissage conçus selon une logique pédagogique cohérente et explicitement réfléchi.

Ce premier aperçu souligne la nécessité d'associer la recherche scientifique à une réflexion politique et pédagogique. La pertinence scientifique, politique et pratique se rejoignent dans l'exigence de trouver de nouvelles formes de suivi empirique et de réflexion collaborative sur les technologies éducatives fondées sur l'IA.

KENNETH HORVATH, professeur en Sciences de l'éducation, Haute école pédagogique de Zurich. Contact: kenneth.horvath@phzh.ch

ANDREA ISABEL FREI, collaboratrice de projet FNS, Haute école pédagogique de Zurich. Contact: andreaisabel.frei@phzh.ch

MARIO STEINBERG, collaborateur de projet FNS, Haute école pédagogique de Zurich. Contact: mario.steinberg@phzh.ch

Bibliographie

- Al-Amoudi, I., & Latsis, J. (2019):** Anormative black boxes. Artificial intelligence and health policy. Dans: I. Al-Amoudi & E. Lazega (éd.), *Post-Human Institutions and Organizations: Confronting the Matrix* (p.119-142). Routledge. <https://doi.org/10.4324/9781351233477>
- Bartolomé, A., Castañeda, L., & Adell, J. (2018):** Personalisation in educational technology: The absence of underlying pedagogies. *International Journal of Educational Technology in Higher Education*, 15(1), 14. <https://doi.org/10.1186/s41239-018-0095-0>
- Beer, D. (2017):** The social power of algorithms. *Information, Communication & Society*, 20(1), 1-13. <https://doi.org/10.1080/1369118X.2016.1216147>
- Dencik, L., Hintz, A., Redden, J., & Tréré, E. (2019):** Exploring Data Justice: Conceptions, Applications and Directions. *Information, Communication & Society*, 22(7), pp. 873-881. <https://doi.org/10.1080/1369118X.2019.1606268>
- Dixon-Román, E., Nichols, T. P., & Nyame-Mensah, A. (2020):** The racializing forces of/in AI educational technologies. *Learning, Media and Technology*, 45(3), 236-250. <https://doi.org/10.1080/17439884.2020.1667825>
- Dourish, P. (2016):** Algorithms and their others: Algorithmic culture in context. *Big Data & Society*, 3(2), 205395171666512. <https://doi.org/10.1177/2053951716665128>
- du Boulay, B. (2019):** Escape from the Skinner Box: The case for contemporary intelligent learning environments. *British Journal of Educational Technology*, 50(6), pp. 2902-2919. <https://doi.org/10.1111/bjjet.12860>
- Francis, P., Broughan, C., Foster, C., & Wilson, C. (2020):** Thinking critically about learning analytics, student outcomes, and equity of attainment. *Assessment & Evaluation in Higher Education*, 45(6), pp. 811-821. <https://doi.org/10.1080/02602938.2019.1691975>
- Gibson, D. (2017):** Big Data in Higher Education: Research Methods and Analytics Supporting the Learning Journey. *Technology, Knowledge and Learning*, 22(3), pp. 237-241. <https://doi.org/10.1007/s10758-017-9331-2>
- Horvath, K. (2020):** Klassifikationen zwischen Politik und Ungleichheit. Pädagogisches Unterscheidungswissen in migrationsgesellschaftlichen Kontexten [Habilitation, Université de Lucerne]. <https://osf.io/preprints/socarxiv/pzjgw/>
- Horvath, K., Steinberg, M., & Frei, A.I. (2023):** Bridging Inquiry and Critique: a neo-pragmatic perspective on the making of educational futures and the role of social research, online first in *Learning, Media and Technology*. <https://doi.org/10.1080/17439884.2023.2193412>
- Jarke, J., & Macgilchrist, F. (2021):** Dashboard stories: How narratives told by predictive analytics re-configure roles, risk and sociality in education. *Big Data & Society*, 8(1), 20539517211025560. <https://doi.org/10.1177/20539517211025561>
- Jananoff, S. (2015):** Future imperfect: Science, technology, and the imaginations of modernity. In *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power*. University of Chicago Press. <https://doi.org/10.7208/chicago/9780226276663.001.0001>
- Kirkwood, A., & Price, L. (2014):** Technology-enhanced learning and teaching in higher education: What is 'enhanced' and how do we know? A critical literature review. *Learning, Media and Technology*, 39(1), pp. -36. <https://doi.org/10.1080/17439884.2013.770404>
- Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014):** Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology*, 106(4), pp. 901-918. <https://doi.org/10.1037/a0037123>
- Neyland, D., & Möllers, N. (2017):** Algorithmic IF ... THEN rules and the conditions and consequences of power. *Information, Communication & Society*, 20(1), pp. 45-62. <https://doi.org/10.1080/1369118X.2016.1156141>
- Noble, S. U. (2018):** Algorithms of oppression: How search engines reinforce racism. New York University Press.
- Perrotta, C., & Selwyn, N. (2020):** Deep learning goes to school: Toward a relational understanding of AI in education. *Learning, Media and Technology*, 45(3), pp. 251-269. <https://doi.org/10.1080/17439884.2020.1686017>
- Prinsloo, P. (2020):** Of 'black boxes' and algorithmic decision-making in (higher) education - A commentary. *Big Data & Society*, 7(1), pp. 1-6. <https://doi.org/10.1177/2053951720933994>
- Riazy, S., Simbeck, K., & Schreck, V. (2020):** Fairness in Learning Analytics: Student At-risk Prediction in Virtual Learning Environments: Proceedings of the 12th International Conference on Computer Supported Education, pp. 15-25. <https://doi.org/10.5220/0009324100150025>
- Selwyn, N. (2022):** The future of AI and education: Some cautionary notes. *European Journal of Education*, 00, pp. 1-12. <https://doi.org/10.1111/ejed.12532>

Skinner, B. F. (1958): Teaching Machines. *Science*, 128(3330), pp. 2969–2977.

Verbert, K., Ochoa, X., De Croon, R., Dourado, R. A., & De Laet, T. (2020): Learning analytics dashboards: The past, the present and the future. *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, pp. 35–40. <https://doi.org/10.1145/3375462.3375504>

Whitman, M. (2020): «We called that a behavior»: The making of institutional data. *Big Data & Society*, 7(1), <https://doi.org/10.1177/2053951720932200>

Williamson, B. (2021): Education Technology Seizes a Pandemic Opening. *Current History*, 120(822), pp. 15–20. <https://doi.org/10.1525/curh.2021.120.822.15>