

Intégration responsable de l'IA Générative dans l'Éducation : proposition d'un plan d'actions stratégiques dirigé par les risques liés aux questions éthiques

Rémi Saurel^{1,2}[0009-0003-7969-1018], Franck Silvestre¹[0000-0002-1134-8200],
Jean-Baptiste Raclet¹[0000-0001-7357-912X] et Emmanuel Lescure²

¹ IRT, Université de Toulouse

² Kosmos Education

Résumé. Les résultats des premières études sur l'intégration de l'Intelligence Artificielle Générative (IAG) dans l'Éducation permettent de dégager 3 tendances pour les pays du *Global North* : les usages se développent très vite et de manière significative, mais avec un décalage croissant des usages en faveur des élèves, qui utilisent l'IAG dans un cadre non-contrôlé par les enseignants et pas toujours par les parents. Dans ce papier de positionnement, nous proposons un cadre d'actions stratégiques issu d'une analyse des risques liées aux questions éthiques identifiées par l'UNESCO. Notre approche vise à limiter les risques à criticité élevée afin d'augmenter les chances d'intégration responsable de l'IAG dans l'Éducation. Nous illustrons notre propos par la description du projet Konsolidation Dashboard (KDo), une initiative de recherche visant à informer et former sur les questions de l'intégration de l'IAG au sein de Tableaux de Bord d'Apprentissage dédiés aux enseignants.

Mots-clé: Intelligence artificielle générative · Éducation · Éthique · Analyse de risques · Littératie numérique

Abstract The results of the first studies on the integration of Generative Artificial Intelligence (GenAI) in education reveal 3 trends for the Global North countries: uses are developing very quickly and significantly, but with a growing shift in the use of GenAI in favour of learners who use GenAI in a context that is not controlled by teachers and not always by parents. In this position paper, we propose a framework for strategic actions based on an analysis of the risks associated with ethical issues identified by UNESCO. Our approach aims to limit high-criticality risks in order to increase the chances of responsibly integrating GenAI into education. We illustrate our position by describing the Konsolidation Dashboard (KDo) project, a research initiative aimed at informing and training on the issues involved in integrating GenAI into Learning Analytics Dashboards for teachers.

Keywords Generative Artificial Intelligence · Education · Ethic · Risk analysis · Digital literacy.

1 Introduction

L'intégration des technologies d'Intelligence Artificielle Générative (IAG) dans l'éducation suscite un intérêt croissant à l'échelle mondiale. Cette adoption rapide reflète les avancées technologiques majeures et leur potentiel à transformer les pratiques éducatives. Cependant, l'introduction de ces outils ne s'est pas faite de manière homogène, en particulier entre les principaux acteurs de la communauté éducative. Les études menées au Royaume-Uni [48] et en France [47] révèlent que les élèves mobilisent ces outils de manière plus importante que les enseignants, avec des disparités selon les niveaux d'éducation (secondaire vs. supérieur) et les contextes géographiques. Ces différences d'adoption soulèvent des questions quant à la manière dont ces technologies s'intègrent dans des cadres éducatifs variés.

Cependant, les analyses actuelles sur le sujet présentent des limites importantes. Les études disponibles se concentrent principalement sur le supérieur et sur un outil particulier (ChatGPT) [54, 8]. Elles manquent également d'un cadre théorique pour analyser les politiques éducatives et se focalisent quasi exclusivement sur les pays du « Global North » — des régions développées comme l'Europe, l'Amérique du Nord et quelques pays d'Océanie [35, 21]. Cette concentration géographique limite la compréhension de l'impact global de l'IAG, tandis que l'évolution rapide des technologies et des usages invite à la prudence quant à l'interprétation des résultats, y compris ceux des études récentes.

Même avec ces limites, toutes les recherches que nous avons identifiées convergent sur certains aspects comme l'évolution rapide des technologies, des usages et du décalage d'utilisation entre élèves et enseignants. L'intégration de l'IAG, notamment du côté des élèves, se fait souvent de manière non pertinente pour l'amélioration des apprentissages. Du côté des enseignants, l'engagement plus modéré est attribué à un manque de compétences et de connaissances sur ces outils et leurs usages éducatifs.

Les résultats des premières études ne permettent pas de formuler des constats généralisables, mais sont suffisamment alarmants pour nous inviter à proposer des stratégies d'intégration de l'IAG dans l'éducation, guidées par une analyse de risques.

Nous nous intéressons dans ce papier de positionnement à la manière de limiter les risques à criticité élevée liés à l'intégration de l'IAG dans l'éducation.

La suite de cet article est composée de 4 sections. La section 2 présente les résultats révélés par les quelques études réalisées sur l'intégration de l'IAG dans l'éducation. La section 3 propose une analyse des risques s'appuyant sur les constats émergents des études à notre disposition et sur les dimensions éthiques documentées par l'UNESCO [50]. La section 4 définit un plan de gestion des risques incluant l'ensemble des parties prenantes du secteur éducatif. Enfin, la section 5 présente le projet *KDo* (Konsolidation Dashboard), un exemple concret d'initiative de recherche servant à informer et former sur les questions de l'intégration de l'IAG au sein de TBA (Tableaux de Bord d'Apprentissage) dédiés aux enseignants. La conclusion du papier insiste sur le caractère non définitif de

la proposition inhérent au contexte d'évolutions rapides et aux incertitudes induites.

2 Intégration de l'IAG dans l'éducation

L'intégration de l'IAG dans l'éducation est un phénomène en pleine expansion, avec des implications importantes pour les acteurs du secteur. Les études actuelles présentent un certain nombre de limites mentionnées dans l'introduction. Ces limites sont catalysées par les évolutions rapides de la technologie, des usages autour de l'IAG et de son intégration d'un point de vue pédagogique. C'est pour cela que nous identifions deux aspects majeurs qui guideront cette section : la rapidité d'évolution technologique et la rapidité d'évolution des usages.

2.1 Une rapidité d'évolution technologique

L'arrivée de ChatGPT en novembre 2022 a bouleversé un ensemble de secteurs, dont l'éducation. Tandis que les premiers grands modèles de langage (*Large Language Model*, *LLM*) étaient principalement connus par la communauté de l'IA, ChatGPT (avec son modèle GPT-3.5 lancé en 2022) a démocratisé l'accès et l'utilisation de ces technologies par le grand public. La puissance de calcul utilisée pour pré-entraîner ces modèles a grandement augmenté, permettant des avancées significatives en termes de qualité des réponses générées. Cette rapidité d'évolution technologique est présente sur deux niveaux : la puissance de calcul et la diversité de LLM.

La puissance de calcul. L'un des moteurs principaux de l'accélération technologique dans le domaine des modèles de langage réside dans l'augmentation de la puissance de calcul disponible. Cette évolution a permis de pré-entraîner des modèles de plus en plus complexes et performants.

Comme le souligne l'UNESCO dans son rapport [50], en seulement trois ans, les modèles sont passés de 117 millions de paramètres (pour GPT-1 en 2019) à 175 milliards de paramètres (pour GPT-3 en 2020). Ce bond massif dans le nombre de paramètres illustre l'augmentation des capacités des modèles à générer des réponses précises, à capter des nuances linguistiques et à étendre leur compréhension contextuelle.

La quantité de données utilisées pour l'entraînement a également connu une expansion sans précédent. Par exemple, GPT-1 a été formé sur un ensemble de 40 Go de données, tandis que GPT-3 utilisait 17 000 Go. En ce qui concerne GPT-4, bien que les chiffres exacts n'aient pas été publiés officiellement, l'UNESCO mentionne que ce modèle aurait été entraîné sur des données atteignant 1 000 000 Go, avec 170 000 milliards de paramètres.

La diversité de LLM. Tandis que les premiers modèles de langage, tels que GPT-3, étaient généralistes, de plus en plus de modèles sont conçus pour des cas d'utilisation spécifiques.

OpenAI, par exemple, a développé une gamme variée de modèles adaptés à différents usages, allant de la génération de texte (GPT) à la capacité de réflexion poussée (o1) [19]. META a introduit LLaMA (Large Language Model Meta AI) [13] avec différents modèles plus ou moins exigeants en ressources, tandis que des entreprises comme Mistral se spécialisent dans des solutions plus ciblées, telles que Pixtral (modèle multimodal) [1] et Ministral (un modèle compact pour des appareils à faible puissance de calcul) [32]. Midjourney, de son côté, est devenu une référence dans la création d'images, avec des versions comme Midjourney V6 et Niji Model 6 qui répondent à des besoins plus spécifiques, tels que la création artistique [31].

Par ailleurs, ces modèles diffèrent également sur leur nombre de paramètres (ou poids). Certains, comme GPT-4, comportent des centaines de milliards de paramètres et sont destinés à une utilisation intensive par des organisations dotées d'une infrastructure puissante. D'autres, plus petits, visent à rendre l'IA accessible sur des appareils individuels ou dans des environnements avec des ressources limitées.

2.2 Une rapidité d'évolution des usages

Tandis que les modèles de langage évoluent rapidement, les usages de ces technologies se multiplient, en particulier dans le domaine de l'éducation.

Malgré une adoption rapide de l'IAG par les apprenants mais aussi les enseignants, le décalage d'usage entre ces deux acteurs est significatif. Selon le *National Literacy Trust* [36], au Royaume-Uni, le pourcentage de 13-18 ans utilisant des outils d'IAG est passé de 37% en 2023 à 77% en 2024. Du côté des enseignants, toujours au Royaume-Uni, cette évolution a été beaucoup moins importante passant de 31% à (seulement) 48%. Cet écart grandissant entre les élèves et les enseignants est également observé en France d'après le rapport du Sénat de 2024-2025 sur le thème de l'IA et Éducation [47].

À partir de ces éléments, nous identifions deux tendances émergentes au sein des pays du *Global North* :

Tendance émergente 1

Les usages se développent très vite et de manière significative, en particulier chez les élèves, mais aussi chez les enseignants.

Tendance émergente 2

Il existe un décalage croissant de l'usage de l'IAG entre les élèves et les enseignants en faveur des élèves.

D'après l'étude du gouvernement du Royaume-Uni, parmi les élèves utilisant l'IAG dans un contexte éducatif, celle-ci est principalement réservée aux devoirs (61%) et à l'apprentissage à la maison (40%) [48]. Cette utilisation en dehors du cadre scolaire, non-contrôlée par les enseignants, soulève des inquiétudes quant à la supervision de ces usages. De plus, toujours d'après cette étude, 30% des

parents d'élèves (de 7 à 13 ans) au Royaume-Uni ne sont pas au courant de l'utilisation de l'IAG par leurs enfants [48].

Tendance émergente 3

Les élèves utilisent l'IAG dans un cadre non-contrôlé par les enseignants et pas toujours par les parents.

Dans la suite du document, nous considérons que l'intégration *responsable* d'une technologie dans l'Éducation est une intégration qui fait progresser de manière probante les apprentissages tout en prenant en compte les questions éthiques posées par cette intégration.

Les trois tendances identifiées nous suggèrent qu'une intégration responsable de l'IAG dans l'Éducation ne va pas de soi : ni les enseignants, ni les élèves ne paraissent armés des connaissances et compétences requises. Nous estimons qu'un plan stratégique d'urgence doit être mis en œuvre pour augmenter les chances d'une intégration responsable de l'IAG dans l'Éducation. Face au degré d'incertitude induit par la rapidité des évolutions tant sur le volet technique que sur le volet des usages, nous proposons de définir ce plan stratégique dans une démarche dirigée par les risques.

3 Des risques de l'intégration de l'IAG dans l'éducation

Dans cette section, en nous inscrivant dans les processus de management des risques décrits dans le PMBOK (Project Management Body of Knowledge) [18], nous identifions et structurons dans un premier temps un ensemble de risques liés à l'intégration de l'IAG dans l'éducation en nous appuyant sur le rapport de l'UNESCO [50]. En suivant, nous proposons une analyse qualitative de ces risques afin d'en estimer la criticité en combinant l'évaluation de leur impact et leur probabilité d'occurrence.

Notre analyse n'englobe pas tous les risques potentiels associés à l'IAG, car la rapidité d'évolution technologique et des usages fait apparaître en permanence de nouveaux défis. Slattery et al. [45] ont travaillé sur le MIT AI Risk Repository, un projet qui vise à identifier et à mettre à jour régulièrement les nouveaux risques afin d'aider chercheurs, développeurs et les différentes parties prenantes dans le domaine de l'IA. À partir de leur analyse, plus de 1000 risques ont été identifiés à l'aide de 50 cadres de risques liés à l'IA. Ils proposent 2 taxonomies : une taxonomie causale (comment, quand et pourquoi ces risques se produisent) et une taxonomie de domaine (7 domaines et 23 sous-domaines) pour classer ces risques.

Sur la base de cet exemple, nous sommes conscients que notre analyse n'est pas exhaustive et que d'autres risques sont déjà présents, tels que ceux concernant la dimension environnementale, ou pourraient émerger à l'avenir. Par conséquent, l'objectif de ce document n'est pas d'être exhaustif, mais plutôt d'initier une approche pour traiter ces risques en évaluant leur criticité, et d'illustrer cette

approche sur un sous-ensemble de risques balisé par le rapport de l'UNESCO [50].

3.1 Identification des risques

À partir du rapport de l'UNESCO [50, p.38] et de ses 5 dimensions éthiques liées exclusivement à l'utilisation des technologies d'IAG dans l'éducation et des descriptions de ces dimensions, nous avons extrait 8 risques, notés de (R1) à (R5.2) dans le Tableau 1.

Parmi ces risques, nous pouvons noter l'aggravation des inégalités éducatives (R1) liée à la dimension éthique *Accès et équité*, induite par le risque d'accroissement des disparités dans l'accès aux technologies et ressources éducatives.

La dimension *Connexion humaine* est quant à elle menacée par la réduction des interactions sociales et des aspects socio-émotionnels (R2) avec l'usage intensif d'agents artificiels interagissant en langage naturel.

Les risques liés au développement intellectuel humain sont multiples, allant de la limitation de l'autonomie et de la capacité d'agir (R3.1) à l'affaiblissement du développement intellectuel à long terme (R3.2) induit notamment par la fourniture de solutions préconçues par l'IAG.

La dimension *Impact psychologique* soulève des inquiétudes quant à la dégradation du bien-être émotionnel (R4.1) en lien direct avec la réduction des aspects socio-émotionnels (R2). Le risque de manipulation et d'influence malveillante (R4.2) a déjà fait l'objet d'une étude portant sur les interactions entre des enfants et des agents d'IA [12] comme Alexa d'Amazon ou Google Home.

Enfin, la dimension *Préjugé et discrimination cachée* soulève des inquiétudes quant à la génération de préjugés (R5.1) et de nouvelles formes de discriminations avec les résultats incertains des modèles d'IAG avec les phénomènes de biais et d'hallucinations [30, 20].

3.2 Estimation de l'impact

D'un point de vue analyse de risque, nous associons à chacun des risques du Tableau 1 un impact fort, car ils conduisent tous à une altérité des caractéristiques importantes de l'être humain (interactions sociales, capacités intellectuelles, bien-être, etc.).

Afin d'évaluer la criticité de ces risques, nous nous intéressons dans la section suivante à l'estimation de leur probabilité d'occurrence.

3.3 Estimation de la probabilité d'occurrence

Notre démarche pour fournir une estimation de la probabilité d'occurrence a consisté à relever, dans les études à notre disposition, les occurrences avérées de ces risques. Nous avons estimé que la présence dans les premières études de certains risques était suffisante pour affecter une probabilité d'occurrence élevée à ces derniers. Notre analyse a porté sur les modalités d'usages par les élèves et les enseignants.

| Dimension éthique | ID | Risque |
|-----------------------------------|------|---|
| Accès et équité | R1 | Aggravation des inégalités éducatives |
| Connexion humaine | R2 | Réduction des interactions entre humains et des aspects socio-émotionnels |
| Développement intellectuel humain | R3.1 | Limitation de l'autonomie et de la capacité d'agir |
| | R3.2 | À long terme, affaiblissement du développement intellectuel |
| Impact psychologique | R4.1 | Dégradation du bien-être émotionnel |
| | R4.2 | Manipulation et influence malveillante |
| Préjugé et discrimination cachée | R5.1 | Génération de préjugés |
| | R5.2 | Génération de nouvelles formes de discriminations |

Tableau 1. 5 Dimensions éthiques et risques associés

Élèves. L'utilisation des outils d'IAG par les élèves est en forte croissance. Selon une enquête réalisée au Royaume-Uni, le pourcentage d'élèves entre 13 et 18 ans utilisant ces outils est passé de 37 à 77% entre 2023 et 2024 [36]. Cette adoption rapide se concentre principalement sur des activités dans un contexte hors-classe, notamment les devoirs à la maison. Ce contexte soulève des inquiétudes quant au contrôle et à la supervision de ces usages [16], renforcées par le fait que les parents ne sont pas toujours conscients de l'utilisation de ces outils par leurs enfants [48]. En lien avec la dimension *Accès et équité* et le risque R1, cette utilisation à la maison peut accentuer les inégalités entre les élèves, en fonction de leur accès à ces technologies. L'utilisation massive peut être associée à la dimension *Connexion humaine* et le risque R2 en raison de la réduction des interactions sociales et de l'impact sur le développement cognitif des élèves.

Une étude menée par Kabir et al. [23] sur l'utilisation d'IAG pour résoudre des questions de programmation informatique a mis en évidence une tendance frappante : malgré la capacité de discerner les réponses de ChatGPT, 35% des élèves préfèrent les réponses générées par l'IA. Cependant, parmi ces réponses, trois sur quatre étaient incorrectes. Des résultats en lien avec cette étude ont été proposés par Darvishi et al. [9] lors d'un essai contrôlé randomisé. Celle-ci montre que les élèves tendent à s'appuyer sur l'assistance de l'IAG plutôt que d'apprendre à partir de celle-ci. Ces situations révèlent et renforcent le manque de compétences critiques telles que l'analyse, la synthèse et l'autonomie dans l'apprentissage, en lien direct avec R3.1 sur la dimension *Développement intellectuel humain*.

L'étude de Druga et al. [12] montre que les enfants personnifient et traitent les agents comme des humains, les perçoivent comme « amicaux » et « de confiance

». Cette imitation de l'interaction peut avoir des effets psychologiques négatifs comme la manipulation et une influence malveillante, identifiée par le risque R4.2 de la dimension *Impact psychologique*.

Enfin, des *Préjugés et discriminations cachées* (dimension des risques R5.1 et R5.2) peuvent être exacerbés par les biais et les hallucinations générés par les modèles d'IAG [30, 20], et ces effets sont particulièrement marqués si les élèves ne sont pas formés à les détecter.

Ces risques deviennent encore plus préoccupants lorsque les élèves utilisent ces outils en dehors du cadre scolaire, sans supervision des enseignants et parfois sans que les parents en soient informés. Les enseignants ont donc un rôle crucial à jouer pour accompagner les élèves dans l'utilisation de ces technologies.

Enseignants. Selon le *National Literacy Trust*, le pourcentage d'enseignants du Royaume-Uni utilisant l'IAG est passé de 31% à 47% entre 2023 et 2024 [36]. Malgré cette augmentation, elle reste largement inférieure à celle des élèves avec plus de 77% d'utilisation pour les 13-18 ans, toujours d'après les résultats du *National Literacy Trust*.

Le DfE (Department for Education) du Royaume-Uni a réalisé un sondage en avril 2023 auprès de 2034 enseignants et 502 chefs d'établissement du pays sur l'utilisation de l'IAG. 83% des enseignants et chefs d'établissement ont indiqué ne pas avoir utilisé d'IAG, représentant environ 2100 personnes. Les résultats montrent que pour 58% de cet échantillon, la principale raison est leur manque de connaissance sur l'IAG et les cas d'usages pédagogiques possibles. Autrement dit, les enseignants et chefs d'établissement ne disposent pas des connaissances et compétences nécessaires pour intégrer l'IAG dans leur enseignement.

La littératie numérique regroupe un ensemble de compétences requises par les individus pour utiliser les technologies de manière efficace et critique. Elle inclut la capacité à rechercher, évaluer, utiliser et créer de l'information en utilisant les technologies numériques. Dans le contexte de l'éducation, la littératie numérique est essentielle pour permettre aux enseignants de guider les apprenants dans l'utilisation des outils numériques de façon responsable [43].

En 2023, seulement sept pays (Chine, Espagne, Finlande, Géorgie, Qatar, Thaïlande et Turquie) avaient développé des ressources spécifiques pour accompagner les enseignants dans l'intégration de l'IA dans leurs pratiques pédagogiques [50]. Cela souligne un retard global dans la formation des enseignants en littératie en IA, malgré une augmentation rapide de l'utilisation de l'IAG par les élèves.

Des initiatives émergent néanmoins. Par exemple, l'UNESCO [52] ou le gouvernement du Royaume-Uni [49] proposent des guides pratiques. En France, ces initiatives proviennent d'universités (Université de Bordeaux, [53]), de laboratoire de recherche (INRIA) avec des formations dédiées [17], et des communautés de réflexion [11] pilotées par la DNE (Direction du Numérique pour l'Éducation). Ces efforts visent à fournir aux enseignants les connaissances et compétences nécessaires pour intégrer l'IAG dans leurs méthodologies pédagogiques. Cependant, ces ressources semblent insuffisantes. En France, dans un

rapport du Sénat, une illustration de cette lacune est la prise de position des syndicats : « *L'ensemble des syndicats insistent sur les besoins de formation très importants des enseignants et la relative impréparation des établissements scolaires* » [47].

En synthèse, les études à notre disposition font état d'évènements qui sont des occurrences directes des risques R2, R3.1 et R4.2. Les enseignants manquant actuellement, pour beaucoup d'entre eux, de littératie en IAG, ne peuvent contribuer que faiblement à la prévention de ces risques.

3.4 Une criticité élevée pour certains risques

En analyse de risques, une échelle à 4 niveaux est souvent utilisée pour évaluer l'impact (de faible à très important) et la probabilité d'occurrence (de très improbable à très probable). La criticité, combinant l'impact et la probabilité d'occurrence peut varier de très basse à très élevée. Nous synthétisons les évaluations des risques liés à l'intégration de l'IAG dans l'Éducation dans le Tableau 2.

| ID | Libellé | Impact | Probabilité | Criticité |
|------|---|----------------|---------------|-------------|
| R1 | Aggravation des inégalités éducatives | Très important | ? | ? |
| R2 | Réduction des interactions entre humains et des aspects socio-émotionnels | | Très probable | Très élevée |
| R3.1 | Limitation de l'autonomie et de la capacité d'agir | | Très probable | Très élevée |
| R3.2 | À long terme, affaiblissement du développement intellectuel | | ? | ? |
| R4.1 | Dégradation du bien-être émotionnel | | ? | ? |
| R4.2 | Manipulation et influence malveillante | | Très probable | Très élevée |
| R5.1 | Génération de préjugés | | ? | ? |
| R5.2 | Génération de nouvelles formes de discriminations | | ? | ? |

Tableau 2. Impact, probabilité et criticité des risques

Dans les études identifiées, un manque de données probantes ne nous permet pas de déterminer la criticité pour 5 des risques identifiés. Cependant, les risques pour lesquels nous avons des éléments nous incitent à les associer à une criticité élevée. Nous proposons un plan de gestion se focalisant sur des actions stratégiques visant à limiter les risques à criticité élevée.

4 Proposition d'un plan de gestion de risques

4.1 Implication de l'ensemble des parties prenantes

La littérature sur les Environnements Informatiques pour l'Apprentissage Humain (EIAH) souligne l'importance d'une approche centrée sur l'humain [37, 14, 3, 56, 29]. Cette approche met en avant la nécessité de considérer les besoins et les attentes des utilisateurs finaux, en l'occurrence les enseignants et les élèves, mais aussi des autres parties prenantes telles que les parents, les organisations en charge de la définition et de la mise en œuvre des politiques éducatives, les chercheurs, les développeurs et les éditeurs de logiciels dans le processus de conception et de déploiement des systèmes.

À travers une revue systématique, Alfredo et al. [3] ont proposé trois recommandations en liant les approches centrées humains et l'IA dans l'éducation:

- l'équilibre sur l'implication des parties prenantes dans la conception et le déploiement des systèmes ;
- l'implication des utilisateurs finaux pour définir l'équilibre entre le contrôle humain et l'automatisation ;
- la conception guidée par la sécurité et la fiabilité.

L'écosystème de l'éducation est complexe et nécessite une collaboration active entre les différents acteurs pour garantir une intégration pertinente de l'IAG dans les pratiques pédagogiques. Défini autour de 4 mots clés - soutenir, fournir, encadrer et informer - nous proposons un schéma de l'écosystème de l'éducation (voir Figure 1) qui met en lumière les interactions entre les différentes parties prenantes.

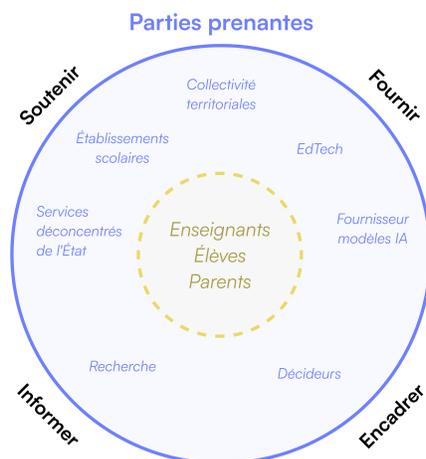


Fig. 1. Parties prenantes et leurs rôles dans l'écosystème de l'éducation

Le soutien, par l'intermédiaire des établissements scolaires et des services déconcentrés de l'État, comprend le suivi, l'accompagnement et la formation des enseignants, des élèves et des parents. Des EdTech (ici, les entreprises des technologies éducatives) et fournisseurs de modèles d'IA doivent fournir services et outils en cohérence avec le contexte (niveau et discipline). L'encadrement consiste à définir les objectifs politiques et le cadre réglementaire, par exemple en produisant des législations et des guides de mise en œuvre comme cela a pu être le cas dans d'autres pays [50]. Enfin, informer l'ensemble des acteurs en mobilisant la recherche et la communication est essentiel pour une prise de décision éclairée.

Wang et al. [54] proposent l'idée que l'effort sur le développement des politiques autour de l'IA doit être partagé entre différentes parties prenantes comme les enseignants, mais aussi les organisations en charge de la définition et de la mise en œuvre des politiques éducatives ou les services déconcentrés de l'État.

La section suivante présente les actions stratégiques que nous proposons pour les parties prenantes jugées prioritaires.

4.2 Actions stratégiques pour cibles prioritaires

Nous identifions 3 cibles prioritaires (Figure 2), enseignants, élèves et parents, pour mener des actions stratégiques en vue d'augmenter les chances d'une intégration responsable de l'IAG dans l'éducation (Figure 1). Les actions au sein de chacun des axes visent la limitation des risques de criticité élevée.

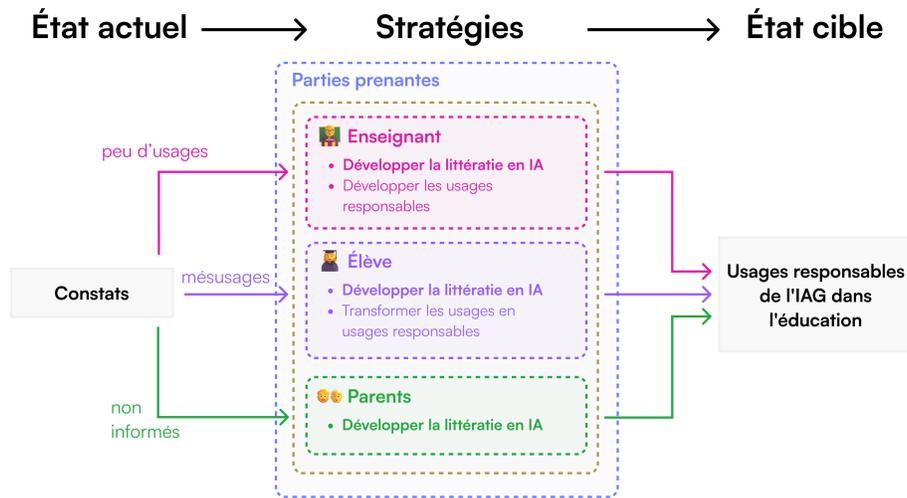


Fig. 2. Axes stratégiques

Le développement de la littératie en IA est une action stratégique commune aux 3 acteurs explicités dans la Figure 2. Long et Magerko [27] ont travaillé sur la définition de la littératie en IA en proposant un ensemble de compétences, tout comme l'UNESCO pour les enseignants [52] et les étudiants [51]. Ng et al. [37] ont conceptualisé cette littératie en utilisant des cadres théoriques issues de la psychologie avec la Taxonomie de Bloom [25] ou encore TPACK (*Technological Pedagogical Content Knowledge*) introduit par Graham [15]). Les résultats de l'étude qualitative de Ding et al. [10] vont dans ce sens : il est important de promouvoir la littératie en IA chez les enseignants pour leur mise en œuvre efficace et responsable des technologies d'IA dans les contextes éducatifs. Cette action de formation sur le développement de la littératie en IA doit permettre aux enseignants de soutenir les élèves dans l'utilisation de l'IAG de manière responsable.

Ainsi, la formation des enseignants à l'IAG n'est pas seulement une question de compétence technique : elle est également un levier stratégique pour préparer les jeunes générations à un monde où l'IA occupe une place centrale. C'est d'ailleurs ce que pense le Sénat français [47] ou les universités internationales [21, 46] : l'IA sera une compétence clé pour les métiers de demain. Un investissement massif dans la formation initiale et continue des enseignants apparaît donc comme une priorité pour tous les systèmes éducatifs.

Van Niekerk et al. [38] ont montré qu'une intervention éducative ciblée modifie significativement les comportements des élèves dans leur utilisation de l'IAG. Après cette formation, les élèves utilisent moins l'IAG pour des éléments cruciaux tels que la structuration, les preuves et les références, et adoptent une approche plus critique vis-à-vis des résultats générés. De plus, ils exploitent mieux les outils pour des activités comme le résumé ou l'adoption d'un ton formel, montrant une meilleure compréhension des forces et des limites de l'IAG. Ces résultats montrent que la formation des élèves à l'IAG est également un levier important pour réduire les risques liés à l'utilisation de ces technologies comme R3.1 (Limitation autonomie et de la capacité à agir) mais aussi R4.2 (Manipulation et influence malveillante) du Tableau 3.1.

Enfin, les parents jouent un rôle crucial dans l'accompagnement des élèves dans l'utilisation de l'IAG en limitant les risques R2 (Réduction des interactions entre humains et des aspects socio-émotionnels) mais aussi R4.2 (Manipulation et influence malveillante). L'utilisation de ces technologies majoritairement dans un contexte hors-classe souligne l'importance de sensibiliser les parents aux enjeux criants en développant une littératie en IA adaptée à leurs besoins.

Ces axes stratégiques ont besoin d'être développés et de s'appuyer sur des données probantes qui manquent cruellement en raison des différences de rythmes entre l'évolution des technologies et celle des pratiques pédagogiques.

Le besoin de données probantes amenées par des recherches est essentiel pour garantir des actions stratégiques efficaces.

4.3 Informer et former

L'effort sur la dimension *Inform*er de la Figure 1 est urgent car on manque de données probantes sur ce qui fonctionne ou non en termes d'apprentissage quand on intègre l'IAG. Il est nécessaire de mener des recherches pour informer et former les acteurs du secteur éducatif sur l'utilisation responsable de l'IAG dans l'éducation. Cependant, cette formation doit se baser sur des travaux relativement récents et avec une rigueur méthodologique élevée pour être pertinente et efficace.

Les travaux de Yan et al. [55] insistent sur trois besoins essentiels pour une intégration réussie de l'IAG dans l'apprentissage humain :

- la littératie en IA;
- la prise de décision basée sur des preuves (ou Evidence-Based decision making);
- la rigueur méthodologique dans la recherche lors de l'utilisation de l'IAG.

Les deux derniers points concernent directement la recherche. Ils sont liés à la nécessité d'établir un cadre théorique robuste pour l'évaluation de la pertinence pédagogique des outils d'IAG. Malgré l'évolution rapide des technologies, la mise en place de cadres et de méthodes d'évaluation rigoureuses des processus, objectifs et expériences pédagogiques, incluant éventuellement des essais contrôlés randomisés à grande échelle, constitue une condition essentielle pour garantir des recherches de qualité permettant de lever une partie des incertitudes entourant l'intégration responsable de l'IAG dans le domaine éducatif.

La section suivante illustre par un exemple de projet de recherche comment celle-ci peut s'insérer dans notre plan de gestion des risques. Les enseignants représentent le premier facteur pouvant limiter les risques liés à l'intégration de l'IAG dans l'éducation. En effet, ils sont les premiers à être en contact avec les élèves et à les accompagner dans l'utilisation de ces technologies dans un cadre éducatif. Il est donc essentiel de démarrer les actions stratégiques envers eux. Le projet KDo s'inscrit dans cette démarche.

5 Le projet KDo

Dans le projet de recherche KDo nous souhaitons étudier comment le numérique et l'IA peuvent soutenir l'implication des enseignants dans la régulation du travail hors classe et pour quels bénéfices.

À travers la recherche et basée sur l'application de travail scolaire hors-classe Konsolidation [39, 40], KDo est un projet qui s'inscrit dans la continuité de la thèse d'Ollivier. En adoptant une approche *Design-Based Research* (DBR) [4, 28], nous plaçons les parties prenantes — enseignants, établissements scolaires et services déconcentrés de l'État (comme les académies) — au cœur du processus de recherche dès le début du projet. Cette implication des parties prenantes est renforcée par le dispositif CIFRE (Convention Industrielle de Formation par

la Recherche), qui permet d'associer étroitement l'EdTech Kosmos³ dès les premières phases de la recherche.

L'étude qualitative d'Ollivier et al. [39, 40] a permis d'explicitier le besoin de certains enseignants de réguler le travail hors-classe à l'aide d'indicateurs actionnables. Ces indicateurs sont issus de l'analyse de l'apprentissage, également connu sous le nom de *Learning Analytics* (LA), qui consiste à collecter et à exploiter les données des apprenants afin de mieux comprendre et optimiser leur processus d'apprentissage [26]. Dans le champ des LA, les tableaux de bord d'apprentissage (TBA) [44] ont déjà montré des résultats intéressants grâce à des indicateurs actionnables [34] (qui permettent d'améliorer l'enseignement à l'aide d'actions clairement identifiables). Cela a été mis en évidence lors d'études sur les types de feedback proposés aux élèves en difficulté [6] ou l'aide à l'identification de ces élèves [42].

Cependant, une conception ou utilisation inadéquate des TBA peut rendre difficile l'analyse, conduire à de fausses conclusions ou être incompatible avec les exigences pédagogiques des enseignants [34, 7, 24]. Pour y remédier, un changement d'orientation est amorcé : des TBA exploratoires vers des TBA explicatifs [14, 41, 22].

Molenaar [33] décrit les interactions enseignant-IA via le concept de *Hybrid Human-AI Regulation* (HHAIR). L'IA, en déchargeant certaines tâches [5], permet aux enseignants de se focaliser sur des activités complexes et à forte valeur ajoutée. Toutefois, l'enseignant doit rester au centre du processus éducatif, en conservant sa « capacité à agir » (*agency*) pour maîtriser les décisions et actions [2].

Dans cette perspective, KDo s'inscrit dans le courant HHAIR en étudiant la mobilisation de l'IAG dans les processus éducatifs. Concrètement, nous souhaitons étudier comment l'IAG pourra être intégrée à l'analyse des productions des élèves afin de fournir des informations actionnables à l'enseignant, tout en permettant une meilleure compréhension des visualisations et des indicateurs pour soutenir une prise de décision éclairée. En offrant une assistance par l'IAG d'abord aux enseignants, le projet KDo s'inscrit dans le premier ensemble d'actions stratégiques : les actions ciblant les enseignants. Le TBA proposé n'a donc pas comme objectif de détecter des usages des IAG fait par les élèves mais vise, in fine, à améliorer les pratiques de révision des élèves notamment au regard des effets tests et espacement.

6 Conclusion

Les résultats des premières études sur l'intégration de l'Intelligence Artificielle Générative (IAG) dans l'Éducation permettent de dégager 3 tendances pour les pays du *Global North* : les usages se développent très vite et de manière significative, mais avec un décalage croissant des usages en faveur des élèves, qui utilisent l'IAG dans un cadre non-contrôlé par les enseignants et pas toujours par les parents.

³ Kosmos Education

À travers une démarche dirigée par les risques et en nous appuyant sur le rapport de l'UNESCO et ses 5 dimensions éthiques, nous avons identifié 8 risques liés à l'intégration de l'IA dans l'éducation. Après une estimation de l'impact et de la probabilité d'occurrences à partir de quelques études disponibles, nous avons identifié une criticité élevée pour certains de ces risques.

Afin de réduire cette criticité, nous avons proposé un plan de gestion des risques orienté vers 3 cibles prioritaires (enseignants, élèves et parents) et impliquant l'ensemble des parties prenantes de l'écosystème de l'éducation.

Le projet KDo illustre comment la recherche peut s'insérer dans ce plan de gestion des risques en s'appuyant sur une approche DBR pour étudier comment le numérique et l'IA peuvent soutenir l'implication des enseignants dans la régulation du travail hors classe et pour quels bénéfices.

References

1. Agrawal, P., Antoniak, S., Hanna, E.B., Bout, B., Chaplot, D., Chudnovsky, J., Costa, D., De Monicault, B., Garg, S., Gervet, T., et al.: Pixtral 12b. arXiv preprint arXiv:2410.07073 (2024)
2. Albert Bandura: Human agency in social cognitive theory (1989), <https://pubmed.ncbi.nlm.nih.gov/2782727/>
3. Alfredo, R., Echeverria, V., Yueqiao, J., Lixiang, Y., Swiecki, Z., Gašević, D., Martinez-Maldonado, R.: Human-centred learning analytics and AI in education: A systematic literature review. *Computers and Education: Artificial Intelligence* **6**, 100215 (Jun 2024). <https://doi.org/10.1016/j.caeai.2024.100215>, <https://www.sciencedirect.com/science/article/pii/S2666920X2400016X>
4. Anderson, T., Shattuck, J.: Design-based research: A decade of progress in education research? *Educational researcher* **41**(1), 16–25 (2012)
5. Buxbaum-Conradi, S., Redlich, T., Branding, J.H.: Conceptualizing Hybrid Human-Machine Systems and Interaction. In: 2016 49th Hawaii International Conference on System Sciences (HICSS). pp. 551–559 (Jan 2016). <https://doi.org/10.1109/HICSS.2016.75>, <https://ieeexplore.ieee.org/document/7427251>, iSSN: 1530-1605
6. Knoop-van Campen, C.A.N., Wise, A., Molenaar, I.: The equalizing effect of teacher dashboards on feedback in K-12 classrooms. *Interactive Learning Environments* **31**(6), 3447–3463 (Aug 2023). <https://doi.org/10.1080/10494820.2021.1931346>, <https://www.tandfonline.com/doi/full/10.1080/10494820.2021.1931346>
7. Campos, F.C., Ahn, J., DiGiacomo, D.K., Nguyen, H., Hays, M.: Making Sense of Sensemaking: Understanding How K–12 Teachers and Coaches React to Visual Analytics. *Journal of Learning Analytics* **8**(3), 60–80 (Jul 2021). <https://doi.org/10.18608/jla.2021.7113>, <https://www.learning-analytics.info/index.php/JLA/article/view/7113>, number: 3
8. Cheng, M.W.T., YIM, I.H.Y.: Examining the use of ChatGPT in public universities in Hong Kong: a case study of restricted access areas. *Discover Education* **3**(1), 1 (Jan 2024). <https://doi.org/10.1007/s44217-023-00081-8>, <https://doi.org/10.1007/s44217-023-00081-8>
9. Darvishi, A., Khosravi, H., Sadiq, S., Gašević, D., Siemens, G.: Impact of AI assistance on student agency. *Computers & Education*

- 210, 104967 (Mar 2024). <https://doi.org/10.1016/j.compedu.2023.104967>, <https://www.sciencedirect.com/science/article/pii/S0360131523002440>
10. Ding, A.C.E., Shi, L., Yang, H., Choi, I.: Enhancing teacher AI literacy and integration through different types of cases in teacher professional development. *Computers and Education Open* **6**, 100178 (Jun 2024). <https://doi.org/10.1016/j.caeo.2024.100178>, <https://www.sciencedirect.com/science/article/pii/S2666557324000193>
 11. DNE-TN2: CREIA – Communauté de Réflexion en Éducation sur l'Intelligence Artificielle (Sep 2024). <https://doi.org/10.58079/12cac>, <https://edunumrech.hypotheses.org/12979>, ISSN: 2725-125X
 12. Druga, S., Williams, R., Breazeal, C., Resnick, M.: "Hey Google is it OK if I eat you?": Initial Explorations in Child-Agent Interaction. In: *Proceedings of the 2017 Conference on Interaction Design and Children*. pp. 595–600. IDC '17, Association for Computing Machinery, New York, NY, USA (Jun 2017). <https://doi.org/10.1145/3078072.3084330>, <https://dl.acm.org/doi/10.1145/3078072.3084330>
 13. Dubey, A., Jauhri, A., Pandey, A., Kadian, A., Al-Dahle, A., Letman, A., Mathur, A., Schelten, A., Yang, A., Fan, A., et al.: The llama 3 herd of models. arXiv preprint arXiv:2407.21783 (2024)
 14. Echeverria, V., Martinez-Maldonado, R., Granda, R., Chiluiza, K., Conati, C., Buckingham Shum, S.: Driving data storytelling from learning design. In: *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*. pp. 131–140. LAK '18, Association for Computing Machinery, New York, NY, USA (Mar 2018). <https://doi.org/10.1145/3170358.3170380>, <https://doi.org/10.1145/3170358.3170380>
 15. Graham, C.R.: Theoretical considerations for understanding technological pedagogical content knowledge (TPACK). *Computers & Education* **57**(3), 1953–1960 (Nov 2011). <https://doi.org/10.1016/j.compedu.2011.04.010>, <https://www.sciencedirect.com/science/article/pii/S0360131511000911>
 16. Hingley, S., Gaskell, E., Morris, J., Man, C., Thornton, A.: Parent, pupil and learner panel 22/23 April/May wave
 17. INRIA: Intelligence artificielle pour et par les enseignants, <https://www.fun-mooc.fr/fr/cours/intelligence-artificielle-pour-et-par-les-enseignants-ai4t/>
 18. Institute, P.M.: A guide to the Project Management Body of Knowledge (PMBOK guide). Project Management Institute (2021)
 19. Jaech, A., Kalai, A., Lerer, A., Richardson, A., El-Kishky, A., Low, A., Helyar, A., Madry, A., Beutel, A., Carney, A., et al.: Openai o1 system card. arXiv preprint arXiv:2412.16720 (2024)
 20. Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y., Chen, D., Dai, W., Chan, H.S., Madotto, A., Fung, P.: Survey of Hallucination in Natural Language Generation (Jul 2024). <https://doi.org/10.48550/arXiv.2202.03629>, <http://arxiv.org/abs/2202.03629>, arXiv:2202.03629
 21. Jin, Y., Yan, L., Echeverria, V., Gašević, D., Martinez-Maldonado, R.: Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. *Computers and Education: Artificial Intelligence* **8**, 100348 (Jun 2025). <https://doi.org/10.1016/j.caeai.2024.100348>, <https://www.sciencedirect.com/science/article/pii/S2666920X24001516>
 22. Jin, Y., Yang, K., Yan, L., Echeverria, V., Zhao, L., Alfredo, R., Milesi, M., Fan, J., Li, X., Gašević, D., Martinez-Maldonado, R.: Chatting with a Learning Analytics Dashboard: The Role of Generative AI Literacy on

- Learner Interaction with Conventional and Scaffolding Chatbots (Nov 2024). <https://doi.org/10.48550/arXiv.2411.15597>, <http://arxiv.org/abs/2411.15597>, arXiv:2411.15597 [cs]
23. Kabir, S., Udo-Imeh, D.N., Kou, B., Zhang, T.: Is Stack Overflow Obsolete? An Empirical Study of the Characteristics of ChatGPT Answers to Stack Overflow Questions (Feb 2024). <https://doi.org/10.48550/arXiv.2308.02312>, <http://arxiv.org/abs/2308.02312>, arXiv:2308.02312
 24. Kaliisa, R., Mørch, A.I., Kluge, A.: My Point of Departure for Analytics is Extreme Skepticism': Implications Derived from An Investigation of University Teachers' Learning Analytics Perspectives and Design Practices. *Technology, Knowledge and Learning* **27**(2), 505–527 (Jun 2022). <https://doi.org/10.1007/s10758-020-09488-w>, <https://doi.org/10.1007/s10758-020-09488-w>
 25. Krathwohl, D.R.: A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice* **41**(4), 212–218 (2002), <https://www.jstor.org/stable/1477405>, publisher: Taylor & Francis, Ltd.
 26. Lang, C., Wise, A., Merceron, A., Gašević, D., Siemens, G.: What is learning analytics. In: Lang, C., Siemens, G., Wise, A.F. (eds.) *The Handbook of Learning Analytics*, pp. 8–18. SOLAR, 2 edn. (2022). <https://doi.org/10.18608/hla22.001>, <https://www.solaresearch.org/publications/hla-22/hla22-chapter1/>
 27. Long, D., Magerko, B.: What is AI Literacy? Competencies and Design Considerations. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. pp. 1–16. CHI '20, Association for Computing Machinery, New York, NY, USA (Apr 2020). <https://doi.org/10.1145/3313831.3376727>, <https://dl.acm.org/doi/10.1145/3313831.3376727>
 28. Mandran, N., Vermeulen, M., Prior, E.: Comment guider les doctorants dans l'utilisation du design-based research? In: *10e Conférence sur les Environnements Informatiques pour l'Apprentissage Humain*. pp. 130–141 (2021)
 29. Matcha, W., Uzir, N.A., Gašević, D., Pardo, A.: A Systematic Review of Empirical Studies on Learning Analytics Dashboards: A Self-Regulated Learning Perspective. *IEEE Transactions on Learning Technologies* **13**(2), 226–245 (Apr 2020). <https://doi.org/10.1109/TLT.2019.2916802>, <https://ieeexplore.ieee.org/document/8713912>, conference Name: IEEE Transactions on Learning Technologies
 30. Maynez, J., Narayan, S., Bohnet, B., McDonald, R.: On Faithfulness and Factuality in Abstractive Summarization (May 2020). <https://doi.org/10.48550/arXiv.2005.00661>, <http://arxiv.org/abs/2005.00661>, arXiv:2005.00661
 31. Midjourney: Midjourney Models, <https://docs.midjourney.com/docs/model-versions>
 32. Mistral: Mistral Models, https://docs.mistral.ai/getting-started/models/models_overview/
 33. Molenaar, I.: The concept of hybrid human-AI regulation: Exemplifying how to support young learners' self-regulated learning. *Computers and Education: Artificial Intelligence* **3**, 100070 (Jan 2022). <https://doi.org/10.1016/j.caeai.2022.100070>, <https://www.sciencedirect.com/science/article/pii/S2666920X2200025X>
 34. Molenaar, I., Knoop-van Campen, C.A.N.: How Teachers Make Dashboard Information Actionable. *IEEE Transactions on Learning Technologies* **12**(3), 347–355 (Jul 2019). <https://doi.org/10.1109/TLT.2018.2851585>, <https://ieeexplore.ieee.org/document/8400475>, conference Name: IEEE Transactions on Learning Technologies

35. Moorhouse, B.L., Yeo, M.A., Wan, Y.: Generative AI tools and assessment: Guidelines of the world's top-ranking universities. *Computers and Education Open* **5**, 100151 (Dec 2023). <https://doi.org/10.1016/j.caeo.2023.100151>, <https://www.sciencedirect.com/science/article/pii/S2666557323000290>
36. National Literacy Trust: Children, young people and teachers' use of generative AI to support literacy in 2024 (Jun 2024), <https://literacytrust.org.uk/research-services/research-reports/children-young-people-and-teachers-use-of-generative-ai-to-support-literacy-in-2024/>
37. Ng, D.T.K., Leung, J.K.L., Chu, S.K.W., Qiao, M.S.: Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence* **2**, 100041 (Jan 2021). <https://doi.org/10.1016/j.caeai.2021.100041>, <https://www.sciencedirect.com/science/article/pii/S2666920X21000357>
38. van Niekerk, J., Delpont, P.M.J., Sutherland, I.: Addressing the use of generative AI in academic writing. *Computers and Education: Artificial Intelligence* **8**, 100342 (Jun 2025). <https://doi.org/10.1016/j.caeai.2024.100342>, <https://www.sciencedirect.com/science/article/pii/S2666920X24001450>
39. Ollivier, D., Silvestre, F., Raclet, J.B., Lescure, E., Broisin, J.: Conception d'un système de révisions: une étude qualitative exploratoire pour identifier les besoins des enseignants et des élèves. In: 11ème Conférence sur les Environnements Informatiques pour l'Apprentissage Humain (EIAH 2023). pp. 36–41 (2023)
40. Ollivier, D., Silvestre, F., Raclet, J.B., Lescure, E., Broisin, J.: Designing a revision system: An exploratory qualitative study to identify the needs of french teachers and students. In: *European Conference on Technology Enhanced Learning*. pp. 294–307. Springer (2023)
41. Pozdniakov, S., Martinez-Maldonado, R., Tsai, Y.S., Echeverria, V., Srivastava, N., Gasevic, D.: How Do Teachers Use Dashboards Enhanced with Data Storytelling Elements According to their Data Visualisation Literacy Skills? In: *LAK23: 13th International Learning Analytics and Knowledge Conference*. pp. 89–99. LAK2023, Association for Computing Machinery, New York, NY, USA (Mar 2023). <https://doi.org/10.1145/3576050.3576063>, <https://doi.org/10.1145/3576050.3576063>
42. Raclet, J.B., Silvestre, F.: Git4school: A dashboard for supporting teacher interventions in software engineering courses. In: *European Conference on Technology Enhanced Learning*. pp. 392–397. Springer (2020)
43. Reddy, P., Sharma, B., Chaudhary, K.: Digital Literacy: A Review of Literature. *International Journal of Technoethics (IJT)* **11**(2), 65–94 (Jul 2020). <https://doi.org/10.4018/IJT.20200701.oa1>, <https://www.igi-global.com/article/digital-literacy/www.igi-global.com/article/digital-literacy/258971>, publisher: IGI Global Scientific Publishing
44. Schwendimann, B.A., Rodríguez-Triana, M.J., Vozniuk, A., Prieto, L.P., Boroujeni, M.S., Holzer, A., Gillet, D., Dillenbourg, P.: Perceiving Learning at a Glance: A Systematic Literature Review of Learning Dashboard Research. *IEEE Transactions on Learning Technologies* **10**(1), 30–41 (Jan 2017). <https://doi.org/10.1109/TLT.2016.2599522>, <https://ieeexplore.ieee.org/document/7542151>, conference Name: IEEE Transactions on Learning Technologies
45. Slattery, P., Saeri, A.K., Grundy, E.A., Graham, J., Noetel, M., Uuk, R., Dao, J., Pour, S., Casper, S., Thompson, N.: The ai risk repository: A comprehensive meta-review, database, and taxonomy of risks from artificial intelligence. *arXiv preprint arXiv:2408.12622* (2024)

46. Sullivan, M., Kelly, A., McLaughlan, P.: ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning & Teaching* **6** (Mar 2023). <https://doi.org/10.37074/jalt.2023.6.1.17>
47. Sénat, Gouvernement Français: IA et éducation, <https://www.senat.fr/rap/r24-101/r24-101.html>
48. UK Government [Social Research]: Generative AI in education: educator and expert views (2024), <https://www.gov.uk/government/publications/generative-ai-in-education-educator-and-expert-views>
49. UK Government [Social Research]: Use Cases for Generative AI in Education: User Research Report (Aug 2024), <https://shorturl.at/OFCoU>
50. UNESCO: Guidance for generative AI in education and research. UNESCO (2023), <https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>
51. UNESCO: AI competency framework for students (2024), <https://www.unesco.org/en/articles/ai-competency-framework-students>
52. UNESCO: AI competency framework for teachers (2024), <https://www.unesco.org/en/articles/ai-competency-framework-teachers>
53. Université de Bordeaux: Activités de pédagogie active pour faire utiliser l'IA (2024), <https://enseigner.u-bordeaux.fr/application/files/5617/0773/6906/Activites-pedagogie-active-IA.pdf>
54. Wang, H., Dang, A., Wu, Z., Mac, S.: Generative AI in higher education: Seeing ChatGPT through universities' policies, resources, and guidelines. *Computers and Education: Artificial Intelligence* **7**, 100326 (Dec 2024). <https://doi.org/10.1016/j.caeai.2024.100326>, <https://www.sciencedirect.com/science/article/pii/S2666920X24001292>
55. Yan, L., Greiff, S., Teuber, Z., Gašević, D.: Promises and challenges of generative artificial intelligence for human learning. *Nature Human Behaviour* **8**(10), 1839–1850 (Oct 2024). <https://doi.org/10.1038/s41562-024-02004-5>, <https://www.nature.com/articles/s41562-024-02004-5>, publisher: Nature Publishing Group
56. Yang, S.J.H., Ogata, H., Matsui, T., Chen, N.S.: Human-centered artificial intelligence in education: Seeing the invisible through the visible. *Computers and Education: Artificial Intelligence* **2**, 100008 (Jan 2021). <https://doi.org/10.1016/j.caeai.2021.100008>, <https://www.sciencedirect.com/science/article/pii/S2666920X21000023>