

LEARNING ANALYTICS : ENJEUX ÉTHIQUES

L'éthique interroge le sens du recours aux Learning Analytics dans la perspective d'une réflexion morale reposant sur le respect de soi, des autres et de ce qui nous entoure. L'éthique implique également la recherche d'un ensemble de principes et de valeurs qui sont à la base d'une sagesse de l'action pour pratiquer les Learning Analytics. La réflexion éthique doit donc mobiliser à la fois une éthique de conviction fondée sur des principes intangibles, sur l'affirmation de valeurs qui donnent un sens à l'action, et une éthique de responsabilité qui s'interroge sur les fins, les moyens et les conséquences des décisions et des actes. Ces deux éthiques s'interpellent pour s'enrichir mutuellement et concernent tant le caractère éthique de la production des données et des analyses, que le caractère éthique de leur utilisation.

Les préoccupations éthiques doivent conduire à examiner les enjeux de la recherche en Learning Analytics en matière de santé publique [1], de respect de la vie privée [2] et d'impacts sociétaux [3].

Enjeux de santé publique

En matière de santé publique, le premier objectif à assurer, à l'occasion d'une recherche en Learning Analytics, consiste à ne pas mettre les participants dans des situations où ils risqueraient de subir des préjudices physiques. À cet égard, les hypothèses de risques paraissent limitées. On peut songer à la question de l'exposition de personnes sensibles (notamment des jeunes enfants) aux radiofréquences. L'hypothèse paraît marginale lorsque, comme c'est le cas le plus fréquent, les recherches en Learning Analytics se focalisent uniquement sur les traces laissées lors de l'interaction avec un environnement type « LMS » (*Learning Management System*, ou plateforme d'apprentissage en ligne, telle Moodle) ou des plateformes de Moocs, etc. Mais elle peut être posée lorsque les recherches utilisent des instruments ou des capteurs tels que des objets connectés. L'Éducation nationale dispose déjà d'un ensemble de préconisations et de bonnes pratiques concernant l'usage du Wi-Fi dans les établissements et les écoles. Et, de manière générale, les risques ne diffèrent pas des autres utilisations (familiales, ludiques, etc.) des objets connectés du quotidien (téléphone portable, tablette, ordinateur, etc.). Lorsque, de manière plus exceptionnelle, la recherche implique des outils se situant près de la tête de jeunes enfants (serre-

têtes pour mesurer les signes d'activité cérébrale, par exemple), il convient sans doute, par mesure de précaution, de minimiser le temps d'exposition à ces outils puisqu'on ne dispose pas d'étude clinique significative sur ce type d'expériences.

Par ailleurs, certains équipements susceptibles d'être utilisés pour des techniques de suivi oculaire (*eye-tracking*) ou des expériences de simulations immersives (réalité augmentée, réalité virtuelle) pourraient éventuellement entraîner des problèmes de « conflit accommodation-vergence » lorsque le système visuel n'a pas achevé son développement. Faute d'études cliniques suffisantes, il convient là encore simplement de faire une place au principe de précaution en présence de jeunes enfants. Par mesure de prudence, l'utilisation d'un matériel adapté à l'enfant et la limitation du temps d'exposition sont alors recommandées.

Enjeux liés au respect de la vie privée

L'analytique des apprentissages soulève la question de la protection des données personnelles, elle-même indissociablement liée au respect de la vie privée. Certes, des limites fonctionnelles sont assignées par les prescriptions du RGPD, qui impliquent que les données personnelles soient collectées pour des finalités déterminées, explicites et légitimes et que seules les données strictement nécessaires pour atteindre ces finalités puissent être collectées et traitées. Mais, en matière de Learning Analytics, il est difficile de déterminer à l'avance quelles informations sur l'environnement et parmi les résultats de l'apprentissage seront susceptibles de se révéler probantes ou, au contraire, anecdotiques. La logique de recherche qui sous-tend l'analytique des apprentissages peut sembler difficilement compatible avec la logique de minimisation de la collecte qui prévaut en matière de réglementation des données personnelles.

Quelles sont les données pertinentes à analyser ? S'agit-il uniquement des données enregistrées dans le cadre d'une activité d'apprentissage ? S'agit-il de données plus larges ? Des données scolaires, des données sociales (étudiant boursier, salarié...) ? Quelle place réserver aux données sensibles telles que les données physiologiques ?

À cet égard, au-delà de l'exigence du respect de la réglementation, il est fondamental qu'une réflexion collective sur les conditions de l'éthique de la récolte et de l'exploitation des données s'instaure.

Bien entendu, les bénéfices attendus des Learning Analytics en matière de respect de la personnalité de l'apprenant sont multiples. En associant les plateformes d'apprentissage en ligne à différents capteurs, les enseignants sont en capacité d'analyser toujours plus finement les réactions de leurs élèves, d'évaluer leur degré d'implication et d'engagement (analyse des fréquences de connexion, des types de lectures, des interactions...) et de proposer un accompagnement proche des besoins de chaque apprenant. Les Learning Analytics conduisent donc à espérer qu'un véritable enseignement « sur mesure », au service de l'apprenant, ne soit pas incompatible avec un enseignement de masse.

Le monitoring de l'apprenant peut également être perçu comme un dispositif de surveillance et de contrôle participant d'une éducation « orwellienne ». La « quantification de soi » permise par le développement des Learning Analytics soulève aussi la question du caractère anxiogène d'une recherche de la performance pouvant être instrumentalisée à des fins très diverses, notamment de contrôle social. Les Learning Analytics ne doivent pas transformer l'apprenant en un producteur passif de données, ni lui assigner de se conformer passivement aux recommandations ou prescriptions d'un algorithme.

Enjeux en matière d'impacts sociaux

La production, l'utilisation et la réutilisation des données de Learning Analytics soulèvent la question du déterminisme, de manière plus ou moins intense selon les objectifs poursuivis (analyse descriptive, prédictive ou prescriptive),

et selon la perception que chaque communauté scientifique se fait du produit de la recherche.

Il semble y avoir une hypothèse implicite selon laquelle la connaissance de l'apprenant et la modélisation prédictive permettent d'identifier les apprenants à risques et de personnaliser les interventions afin d'augmenter les chances de réussite, en permettant de repérer plus précocement des troubles du langage, des troubles arithmétiques ou bien encore des troubles de coordination motrice, et en favorisant une remédiation rapide de nature à lutter contre l'échec. Mais le risque de déterminer un peu trop tôt l'avenir d'un apprenant ne peut être négligé. La question d'un futur prédit par le passé est un dilemme éthique classique dans la réflexion sur le Big Data. À ce titre, il est légitime de s'interroger sur la nature des indicateurs qui permettent de favoriser l'analyse de situations de prédispositions favorables ou de situations à risques (*predictive analytics*) et sur l'utilisation pouvant être faite des profils ainsi obtenus.

Par ailleurs, en favorisant un apprentissage adaptatif fondé sur la personnalisation, les Learning Analytics ouvrent la voie à une pédagogie différenciée, facteur d'une bonne intégration éducative et sociale. Pour autant, les Learning Analytics peuvent aussi susciter des questions concernant le risque de discriminations. Comment anticiper des démarches favorables à la différenciation pédagogique tout en luttant contre l'usage d'indicateurs qui peuvent exclure ? Pour éviter les risques liés à une discrimination implicite ou explicite, il importe de s'assurer que les analyses sont effectuées sur des ensembles de données représentatifs. Il importe, surtout, de permettre à l'apprenant (ou celui qui le représente juridiquement) de prouver l'existence de biais et de stéréotypes dans le profilage, de démontrer que des analyses prédictives sont fausses ou incomplètes lorsqu'une décision lui fait grief, et de favoriser l'exercice du droit à l'oubli.