

Learning Analytics, la data pour (per)former

Laurence Kohli, Elisa Roch, Elodie Perring, Cristian Pottier

Etudiant-e-s en ingénierie des médias, 1^{ère} année, HEIG-VD

Quoiqu'il se passe sur l'internet, le réseau reste un puissant outil de collecte de données ! L'usage le plus courant consiste à dresser des profils d'utilisation, pour mieux cibler, et finalement mieux vendre. Mais quid d'une collecte de données utile à l'éducation ?

I. L'ÉVOLUTION DU DIGITAL DANS L'ÉDUCATION

L'utilisation de jeux informatiques pour certains cours, l'installation de tableaux interactifs dans les établissements scolaires et la mise en place de plateformes éducatives sont autant d'outils que l'on assimile aux EdTech (*Educational Technologies*). Selon des experts en bourse, les dépenses dans le domaine de l'EdTech devraient continuer à croître, passant à plus de 400 milliards de dollars en 2025, ils recommandent même d'investir dans des start-ups qui se spécialisent dans l'éducation [1, 2, 3].

Les termes e-learning ou virtualisation de l'enseignement sont apparus avec les premiers LMS (*Learning Management Systems*) au début des années 2000. Avec ces plateformes logicielles, les utilisateurs se voient attribuer un ou plusieurs rôles (prof, étudiant, etc.) et peuvent s'inscrire à des cours. Les enseignants peuvent ainsi mettre leurs supports de cours en ligne (sous forme de documents, de vidéos ou autre), créer des questionnaires en ligne, recevoir des documents de la part des étudiants et leur donner des feedbacks. Au cours de leur évolution, les LMS ont intégré des outils de LA permettant d'analyser des parcours d'étudiants, dans le but de rendre les processus d'apprentissage plus efficaces et adaptés aux élèves [4].

II. LES ORIGINES DES LA

Progressivement, dans le secteur de la recherche, un intérêt se développe pour la collecte et l'analyse des données produites par l'informatisation de l'enseignement. Pour donner suite à cet intérêt apparaissent presque simultanément les LA (*Learning Analytics*) et EDM (*Educational Data Mining*) [5], deux expressions différentes pour exprimer des préoccupations similaires. Autour de chacun des termes se forment différentes communautés, à l'exemple de l'IEDMS (*International Educational Data Mining Society*) [6] et de SOLAR (*Society for Learning Analytics Research*) à qui l'on doit la première conférence sur les LA, en 2011 (LAK'11¹).

Ces deux entités ont été créées afin de promouvoir une réflexion critique sur l'analyse de l'apprentissage.

En 2012, les plates-formes d'apprentissage en ligne telles que Blackboard et Desire2Learn sont les premières à intégrer des fonctionnalités d'analyse de l'apprentissage, présentées comme utiles à l'éducation moderne.

Puis en 2015[7], Moodle qui compte parmi les LMS les plus

répandus, propose une extension LA durant le *MoodleMoot US*, qui réunit chaque année une large communauté de passionnés du célèbre logiciel.

III. COLLECTER ET ANALYSER LES DONNÉES

On peut distinguer la collecte de données passive, qui concerne par exemple la date de la dernière connexion au cours, l'historique d'activité ou le nombre de ressources ouvertes par l'apprenant, et la collecte active, avec les réponses données à des questionnaires ou des évaluations [8].

Les LA traitent principalement des données de type booléen (vrai ou faux). Les questions de base auxquelles elles peuvent répondre sont : est-ce que la ressource a été ouverte, est-ce que l'apprenant a regardé la vidéo, est-ce que la réponse est correcte. Les évaluations peuvent contenir des questions à choix multiples, des textes à trou, des assemblages de concept, etc. Mais chaque réponse sera juste ou fautive selon les indications données par l'enseignant. Les données ainsi traitées seront visibles par l'enseignant sous forme de tableau de données ou sous une forme synthétique [Fig. 1]. Les étudiants et enseignants peuvent communiquer à travers des feedbacks ou des forums. Certains messages peuvent être automatisés et être envoyés à certains étudiants répondant à des critères prédéfinis. On pourrait par exemple rappeler aux élèves les moins assidus qu'une certaine activité doit être réalisée ou envoyer un encouragement à d'autres [8].

IV. PRÉOCCUPATIONS COURANTES

Ces dernières années [14], les LA suscitent des préoccupations parmi les chercheurs, notamment en raison de son approche exploratoire qui diffère des approches classiques des sciences humaines et sociales (SHS). Certains chercheurs remettent en question la modélisation a posteriori des données d'apprentissage, soulignant l'importance de valider des hypothèses basées sur des modèles théoriques a priori.

De plus [15], l'utilisation croissante des données massives dans l'éducation suscite des interrogations quant à la validité des données, leur capacité à mesurer précisément ce qu'elles sont censées mesurer, et l'effacement du contexte social et biographique des apprenants. Les inquiétudes portent également sur les éventuels biais introduits par les plateformes d'apprentissage en ligne et les interfaces de programmation.

Afin de relever ces défis, il est crucial de développer des approches mixtes combinant des méthodes quantitatives et qualitatives, permettant ainsi de contextualiser les données et de saisir la complexité des processus d'apprentissage.

Malgré la croissance extraordinaire de l'apprentissage en ligne, les utilisations restent trop ciblées et servent uniquement à aider l'apprentissage traditionnel au lieu de devenir une

¹ *Learning Analytics & Knowledge*

nouvelle méthode de travail indépendante et différente [8]. Les institutions devront collaborer et former les enseignants différemment.

La mise en place d'infrastructures adaptées pour la collecte et l'analyse des données devrait devenir une préoccupation majeure. Mais les écoles publiques auront-elles les moyens ?

Enfin, il s'agira de surmonter les résistances et éthiques de la population en misant sur la transparence de l'utilisation des données et la promotion de cette nouvelle technologie.

V. QUESTIONS D'AVENIR

En Europe [9], la conformité au Règlement général sur la protection des données (RGPD) garantit que les données des individus soient préservées lorsqu'ils décident de les partager. Les LA sont ainsi confrontées à un dilemme majeur : comment peuvent-elles préserver en toute sécurité les données des étudiants ? PPTs (*Privacy Preserving Technologies*) [10] ou les technologies préservant la confidentialité sont des techniques cryptographiques qui protègent les données tout en garantissant leur utilisation sécurisée. Les PPTs garantissent une connexion de données plus sécurisée, un partage de données plus étendu et une exploitation plus approfondie des données que les méthodes traditionnelles. Leur adoption est déjà en cours dans certains organismes éducatifs, ce qui témoigne de leur efficacité et de leur potentiel.

Sur un tout autre plan, l'égalité sociale semble menacée. Apple [11] ou d'autres grandes compagnies sont désormais en mesure de fonder des écoles. Assurément, celles-ci seront mieux armées que les écoles publiques pour cibler les besoins des apprenants... grâce aux LA. Selon le montant des frais d'inscription, seul un faible pourcentage de la population pourra y avoir accès, en bénéficiant d'une éducation bien supérieure à la moyenne. Impensable ? Pourtant le débat n'est pas nouveau [12].

Même si une ville d'Irlande interdisait les smartphones au moins de 13 ans suite à une décision collective des parents et des enseignants [13], l'essor de la data pour (per)former ne semble pas prêt de s'épuiser. L'étape suivante consistera-t-elle à entraîner les intelligences artificielles à décerner des diplômes ? Difficile à croire, mais pas complètement surréaliste non plus.

RÉFÉRENCES

- [1] YVES HULMANN, 2022. L'essor de l'«EdTech» n'en est qu'à ses débuts. *Allnews* [en ligne]. 12 mai 2022. Disponible à l'adresse : <https://www.allnews.ch/content/interviews/l%E2%80%99essor-de-l%E2%80%99edtech-n%E2%80%99en-est-qu%E2%80%99%C3%A0-ses-d%C3%A9buts> [consulté le 13 juin 2023].
- [2] Swiss Edtech Collider, [sans date] [en ligne]. Disponible à l'adresse : <https://www.edtech-collider.ch/testbed/> [consulté le 13 juin 2023].
- [3] Solutions Edtech, [sans date] *Intel* [en ligne]. Disponible à l'adresse : <https://www.intel.com/content/www/fr/fr/education/technology/learning-and-teaching.html> [consulté le 13 juin 2023].
- [4] LMS, [sans date] *Digiforma* [en ligne]. Disponible à l'adresse : <https://www.digiforma.com/definition/lms/> [consulté le 13 juin 2023].
- [5] ERIC - EJ1064637 - A Survey of Educational Data-Mining Research, *Research in Higher Education Journal*, 2013-Apr, [sans date] [en ligne]. Disponible à l'adresse : <https://eric.ed.gov/?id=EJ1064637> [consulté le 27 juin 2023].

- [6] International Educational Data Mining Society | UIA Yearbook Profile | Union of International Associations, [sans date] [en ligne]. Disponible à l'adresse : <https://uia.org/s/or/en/1122278867> [consulté le 23 juin 2023].
- [7] Learning Analytics Specification - MoodleDocs, [sans date] [en ligne]. Disponible à l'adresse : https://docs.moodle.org/dev/Learning_Analytics_Specification [consulté le 12 juin 2023].
- [8] LANG, Charles et al. (éd.), 2022. *Handbook of Learning Analytics*. Second. Society for Learning Analytics Research (SoLAR). ISBN 978-0-9952408-3-4.
- [9] Privacy-Preserving Technologies and Rights-Based Privacy Regulation Compliance, 2023 *Luna* [en ligne]. Disponible à l'adresse : <https://www.lunadna.com/privacy-preserving-technologies/> [consulté le 27 juin 2023].
- [10] JACKSON, Mariel, 2022. How can educational institutions safely expand their use of student data? *McCourt School of Public Policy* [en ligne]. 1 août 2022. Disponible à l'adresse : <https://mccourt.georgetown.edu/news/privacy-preserving-technologies-in-education/> [consulté le 23 juin 2023].
- [11] Education Community, [sans date] *Apple Education Community* [en ligne]. Disponible à l'adresse : <https://education.apple.com/en> [consulté le 20 juin 2023].
- [12] Libre choix de l'école: un libre-service inéquitable - *Le Temps*, 2008 [en ligne]. Disponible à l'adresse : <https://www.letemps.ch/opinions/libre-choix-lecole-un-libreservice-inequitable> [consulté le 20 juin 2023].
- [13] Greystones parents agree to 'no smartphone' code for children until second level, [sans date] *The Irish Times* [en ligne]. Disponible à l'adresse : <https://www.irishtimes.com/ireland/education/2023/05/26/greystones-parents-agree-to-no-smartphone-code-for-children-until-second-level/> [consulté le 20 juin 2023].
- [14] Les Learning Analytics en question, [sans date] [en ligne]. Disponible à l'adresse : <https://journals.openedition.org/dms/3485> [consulté le 28 juin 2023].
- [15] CHERIGNY, Florence et al., [sans date]. L'analytique des apprentissages avec le numérique Groupes thématiques de la Direction du numérique pour l'Éducation (DNE -TN2).

ILLUSTRATION

The screenshot shows the Moodle Learning Analytics interface. At the top, it says 'Moodle - Learning Analytics' and the URL is 'http://moodle.org/mod/la.php?id=123'. Below the navigation menu, there is a table titled 'Learning Analytics' with columns for Name, Access, Views, Submission, Status, and Notifications. The table lists four students: Giacomo Gultzoni, Mariah Maclochlan, Valerie Liberty, and Guido Jack Gultzoni. A progress indicator shows 'Ahead (89%)'.

Name	Access	Views	Submission	Status	Notifications
Giacomo Gultzoni	65%	10%	25%	●	2
Mariah Maclochlan	0%	0%	0%	●	1
Valerie Liberty	13%	27%	0%	●	1
Guido Jack Gultzoni	65%	52%	65%	●	1
Guido Jack Gultzoni	96%	88%	78%	●	1

[Fig. 1] Learning Analytics Specification – MoodleDocs