

Vers une modélisation de l'expérience optimale

d'apprentissage via les Learning Analytics

Sergio Iván Ramírez Luelmo¹[0000-0002-7885-0123] – 1ère année de thèse

¹ EA (4354) CIREL, Université de Lille, Cité Scientifique, Lille, 59650, France
sergio.ramirez-luelmo@univ-lille.fr

Abstract. Les Learning Analytics constituent aujourd'hui un sujet d'études fort consolidé dans les champs des Environnements Informatiques pour l'Apprentissage Humain et de l'apprentissage tout au long de la vie. Ce sujet joue un rôle central et médiateur dans l'exploitation, l'analyse, la présentation et l'interprétation des données massives générées par des apprenants de profils hétérogènes dans une variété d'Environnements d'Apprentissage. De plus, le rôle du Modèle d'Apprenant est amplement surligné dans plusieurs études comme facilitateur dans le suivi des apprenants, dans la personnalisation des parcours et contenus, et dans les pratiques numériques des formateurs dans des différents Environnements d'Apprentissage. Ce projet de thèse a pour objet de prendre en compte l'expérience optimale d'apprentissage dans la modélisation de l'apprenant via les Learning Analytics. À cet effet nous souhaitons mettre en œuvre un processus de récolte, de traitement et d'analyse de données numériques d'utilisation d'un EIAH, en vue de modéliser les activités de l'apprenant. Le Modèle d'Apprenant souhaité mobilise les traces d'interaction et les évaluations obtenues par les apprenants utilisateurs de ces Environnements Informatiques pour l'Apprentissage Humain dans un contexte de formation tout au long de la vie.

Keywords: Learning Analytics, Apprentissage Tout-au-long de la Vie, Modèle d'Apprenant, Expérience Optimale d'Apprentissage, Flow.

1 Introduction et contexte

L'utilisation des outils numériques dans la vie de tous les jours, y compris dans les différents contextes éducatifs, donne lieu à une collecte de données d'utilisateurs. Or, le contexte d'apprentissage tout au long de la vie ne nous permet pas de limiter cette collecte aux moyens établis par les institutions éducatives : pratiquement toute interaction en ligne laisse derrière des informations, souvent massives, susceptibles d'être collectées pour produire ce qu'il est convenu d'appeler des traces d'utilisation (Boyer, 2019; Mille, 2013). L'exploitation de ces traces numériques repose sur les techniques associées au Data Mining et aux Data Analytics (Settouti, Prié, Marty, & Mille,

2007). Or, comme dans de nombreux autres domaines¹, le secteur de l'enseignement supérieur reconnaît l'impact positif de ces champs de recherche² sur le développement académique. En effet, les LA visent à fournir aux responsables éducatifs des repères fondés sur le traitement des données, les aidant à améliorer l'efficacité et la qualité de l'enseignement et de l'apprentissage. Parallèlement, les environnements d'apprentissage en ligne massivement multi-apprenants (MOOC) ouvrent de nouvelles perspectives de recherche dans la mesure où la volumétrie des données collectées permet de corroborer *in-situ* la modélisation de l'apprenant.

2 Cadre conceptuel

« On apprend toujours seul, mais jamais sans les autres » (Carré, 2005). Cette formule est d'autant plus en vigueur quand ces échanges se réalisent à l'intérieur d'une communauté virtuelle dans un environnement numérique en ligne (Heutte, 2017), avec des participants à des parcours hétérogènes. Or, des études diverses ont montré les effets positifs de l'expérience optimale³ Flow sur l'apprentissage en ligne, spécifiquement dans un contexte collectif et hétérogène: une meilleure performance, créativité, engagement, motivation, (Jackson & Csikszentmihalyi, 1999), etc. Dans le cadre d'étude des EAIH, l'intérêt à considérer cet état psychologique dans la modélisation des caractéristiques de l'apprenant ne peut pas être ignoré : l'expérience optimale (Flow) permet d'engendrer les éléments nécessaires aboutissant à un temps de travail allongé et à une immersion plus soutenue dans les tâches, améliorant ainsi les résultats des apprenants dans des contextes d'apprentissage en ligne.

Parallèlement, la modélisation de l'apprenant permet la personnalisation des activités d'apprentissage (Mawas, Ghergulescu, Moldovan, Muntean, & Muntean, 2019). Le Modèle d'Apprenant (LM) a comme but d'encoder de manière individuelle les apprenants en se servant d'un ensemble bien défini de dimensions (Nakic, Granic, & Glavinic, 2015) telles que des états cognitifs, des comportements et des préférences d'apprentissage et/ou des préférences personnelles. Or, la complexité de cette modélisation réside dans (1) l'identification et sélection des caractéristiques de l'apprenant qui affectent son apprentissage et, (2) la prise en compte de ses états psychologiques présents pendant cette étape (Abyaa, Khalidi Idrissi, & Bennani, 2019). Une fois identifiées, un processus continu de collecte et mise à jour des données de l'apprenant poursuit en analyse de données et puis, en exploitation des indicateurs obtenus par des utilisateurs humains (Siemens & Baker, 2012). Finalement, l'apprentissage tout au

¹ Ex. les affaires, le commerce électronique ou la santé en ligne.

² Appelées dorénavant Educational Data Mining (EDM) et, Learning Analytics (LA) si c'est dans le cadre d'un contexte éducatif.

³ Le flow (EFRN, 2014) est « un état d'épanouissement lié à une profonde implication et au sentiment d'absorption que les personnes ressentent lorsqu'elles confrontées à des tâches dont les exigences sont élevées et qu'elles perçoivent que leurs compétences leur permettent de relever ces défis ». La notion d'expérience optimale est étroitement liée à celle du flow dans le sens que le flow peut être considéré comme « une expérience optimale au cours de laquelle les personnes sont profondément motivées à persister dans leurs activités ».

long de la vie (ou Lifelong Learning LLL) peut être défini comme l'apprentissage systématique et intentionnel au cours de la vie d'une personne, impliquant les domaines formels et informels (Cropley & Knapper, 1983).

3 Problématique

(Siemens & Baker, 2012) ont cité le rôle des LA de passerelle entre l'informatique et la sociologie/psychologie de l'apprentissage ⁴. L'objectif donc de ce projet de thèse est de contribuer à la réduction des écarts entre les objectifs ciblés à atteindre et les méthodes et moyens à mettre en œuvre pour ce faire (Ferguson, 2012). Plus précisément, entre extraction et collecte des traces et leur analyse et compréhension. C'est-à-dire de répondre à la question de recherche suivante :

Comment modéliser et prédire, à partir des traces numériques massives d'apprentissage, l'expérience optimale d'apprentissage chez les apprenants ?

4 Perspectives

Pour bien mener cet objectif de modélisation, un partenariat de recherche à la fondation I-Site ULNE met en œuvre (1) le partage et l'analyse des corpus de données de traces d'apprenants (MOOC dans un contexte LLL), et (2) des informations et des métadonnées concernant l'environnement et caractéristiques de ces apprenants. Ces corpus comprennent des traces d'utilisation et des résultats auto-rapportés sur leur expérience optimale d'apprentissage. Leurs gestions sont assurées par les responsables chargés de ces plateformes : MOOC « Gestion de Projet » (GdP) à l'École Centrale de Lille, MOOC « Hautes études commerciales » (HEC) à l'Université de Montréal et MOOC « Réseaux-Télécommunications » (RT), à l'Institut Mines Télécom Lille-Douai.

Le planning de travail prévisionnel (12/2019) comprend, (1^{ère}) une revue de littérature en vue de réaliser la synthèse des travaux scientifiques menés sur le sujet, suivie d'une curation afin d'identifier et/ou développer les outils nécessaires pour la collecte des traces. Également, un premier test du protocole conçu sur l'un des terrains partenaires (MOOC GdP) est prévu, ainsi que des communications sur les résultats des recherches effectuées. (2^{ème}) il est prévu d'effectuer des analyses sur les premières données terrain (GdP) et d'en élaborer un premier Modèle suivi d'un test sur un des deux autres terrains identifiés. Un cycle de révision, amélioration et correction du Modèle est mis en place, avec des tests périodiques. Finalement, (3^{ème}) une validation par les pairs du Modèle proposé précèdera le travail de rédaction de la thèse.

⁴ « les domaines techniques, pédagogiques et sociaux doivent être mis en dialogue les uns avec les autres pour assurer que les interventions et les systèmes organisationnels répondent aux besoins de toutes les parties prenantes »

5 Remerciements

Ce projet a été réalisé avec le soutien financier de l'État français dans le cadre du Programme « Investissements d'avenir » (I-SITE ULNE / ANR-16-IDEX-0004 ULNE) géré par l'Agence Nationale de la Recherche française (ANR).

References

- Abyaa, A., Khalidi Idrissi, M., & Bennani, S. (2019). Learner modelling: systematic review of the literature from the last 5 years. *Educational Technology Research and Development*, 67(5), 1105–1143. <https://doi.org/10.1007/s11423-018-09644-1>
- Boyer, A. (2019). Quelques réflexions sur l'exploration des traces d'apprentissage. *Distances et Médiations Des Savoirs*, 27, 0–9. Retrieved from <http://journals.openedition.org/dms/4086>
- Carré, P. (2005). *L'Apprenance*. (Dunod, Ed.). Paris, France.
- Cropley, A. J., & Knapper, C. K. (1983). Higher Education and the Promotion of Lifelong Learning. *Studies in Higher Education*, 8(1), 15–21. <https://doi.org/10.1080/03075078312331379081>
- Ferguson, R. (2012). *The State of Learning Analytics in 2012: A Review and Future Challenges a review and future challenges*. *Media* (Vol. 11). <https://doi.org/10.1504/IJTEL.2012.051816>
- Heutte, J. (2017). L'environnement optimal d'apprentissage : contribution de la recherche empirique sur les déterminants psychologiques de l'expérience positive subjective aux sciences de l'éducation et de la formation des adultes. *Sciences et Bonheur*, 82–99. Retrieved from <https://hal.archives-ouvertes.fr/hal-01597551>
- Jackson, S. A., & Csikszentmihalyi, M. (1999). *Flow in sports: The keys to optimal experiences and performances*. *Flow in sports: The keys to optimal experiences and performances*. Retrieved from <https://psycnet.apa.org/record/1999-02792-000>
- Mawas, N. El, Ghergulescu, I., Moldovan, A.-N., Muntean, C. H., & Muntean, C. H. (2019). *Pedagogical based Learner Model Characteristics*. Retrieved from <http://www.newtonproject.eu/>
- Mille, A. (2013). De la trace à la connaissance à l'ère du Web. Introduction au dossier. *Intellectica. Revue de l'Association Pour La Recherche Cognitive*, 59(1), 7–28. <https://doi.org/10.3406/intel.2013.1083>
- Nakic, J., Granic, A., & Glavinic, V. (2015). *Anatomy of student models in adaptive learning systems: A systematic literature review of individual differences from 2001 to 2013*. *Journal of Educational Computing Research* (Vol. 51). <https://doi.org/10.2190/EC.51.4.e>
- Settouti, L.-S., Prié, Y., Marty, J.-C., & Mille, A. (2007). Vers des Systèmes à Base de Traces modélisées pour les EIAH. *Rapport de Recherche RR-LIRIS-2007*, 1–30. Retrieved from <http://cluster-isle-eiah.liris.cnrs.fr/>
- Siemens, G., & Baker, R. S. J. D. (2012). Learning analytics and educational data mining: Towards communication and collaboration. In *ACM International Conference Proceeding Series* (pp. 252–254). New York, New York, USA: ACM Press. <https://doi.org/10.1145/2330601.2330661>