



D R P I

Direction Recherche
Partenariats Innovation

PROPOSITION DE SUJET POUR UN CONTRAT DOCTORAL

Laboratoire L3i	École doctorale EUCLIDE
Sujet de thèse Intitulé scientifique Recommandation de processus pour une recommandation adaptée dans un outil pédagogique d'accompagnement à l'atteinte d'objectifs Intitulé vulgarisé (<i>explicite pour un non spécialiste</i>) Recommandation de ressources dans un environnement d'apprentissage numérique : apport de la fouille de processus.	
Direction de la thèse <i>Identité du/de la/des directeur-trice-s (grade, HDR) et des éventuels co-encadrant-e-s</i> La Rochelle Université : Ronan Champagnat (MCF-HDR, L3i – 30 %), Mourad Rabah (MCF, L3i – 40 %) Waterford Institute of Technology: Steven Davy (Chercheur, TSSG – 30 %) Collaborateur extérieur : Samuel Nowakowski (MCF-HDR, LORIA)	
Descriptif du sujet <i>Éléments d'explication du sujet (enjeux scientifiques, applicatifs, sociétaux...)</i> Les systèmes d'information pédagogiques se sont rapidement développés avec la mise en place d'environnements numériques de travail. La quantité de données produites et de traces laissées par les utilisateurs de ces systèmes offre l'opportunité de fournir des tableaux de bord et des analyses sur les apprenants. La personnalisation des apprentissages est devenue un facteur essentiel pour la réussite des apprenants. Plusieurs projets se sont développés, dont le projet Nouveaux Coursus Universitaires (NCU) du Programme Investissement d'Avenir (PIA), afin de traiter cette problématique. Ces projets s'attaquent au problème de l'offre de formation et de parcours d'apprentissage personnalisés, avec un accompagnement des étudiants dans leur projet de formation et d'insertion professionnelle. Toutefois peu d'outils sont capables de fournir des indicateurs pertinents afin de recommander des parcours personnalisés ou des actions de remédiation pour les étudiants en difficulté. De leur côté les recherches en Learning Analytics visent à utiliser les techniques de fouille de données appliquées aux données de l'éducation, incluant des informations pour les apprenants et les enseignants, afin d'extraire des informations permettant d'identifier des comportements et de fournir des recommandations (van Harmelen & Workman, 2012). Les objectifs du Learning Analytics concernent la production d'indicateurs pour l'apprenant et pour le formateur, l'adaptation du contenu et la prédiction (Moissa, 2015, cité par (Cadavid & Corcho, 2018)). La découverte des parcours d'apprentissage reste un défi à relever dans le domaine pédagogique. Toutefois, dans des domaines connexes, on trouve des travaux approchants. En particulier des recherches se sont focalisées, à partir de l'analyse des traces d'exécution réelles des utilisateurs (enregistrement des chemins de navigation),	

sur l'extraction de connaissances sur le parcours de l'utilisateur. L'objectif est alors de déterminer pour chaque utilisateur ou par groupe d'utilisateur le comportement adopté lors de l'utilisation du système.

L'équipe eAdapt travaille sur des systèmes adaptatifs pour l'e-Éducation. Les premiers travaux portaient sur la définition d'une architecture et d'heuristiques pour l'adaptation du contenu et des scénarii d'apprentissage (thèse de D. Sawadogo 2016). Nous avons ensuite orienté nos travaux sur l'identification des trajectoires des apprenants en utilisant la fouille de processus. Une première thèse de J. Leblay (2019) nous a permis de démontrer la pertinence d'une telle démarche pour l'extraction des connaissances liées aux trajectoires. La thèse de M. Trabelsi-Hamdi (en cours) porte sur l'identification de trajectoires similaires dans le cadre de l'utilisation d'une bibliothèque numérique. Nous possédons donc la connaissance pour extraire des trajectoires d'apprentissage dans un système pédagogique.

Un système de recommandation trouve une partie de sa justification et de son utilité dans la prise en compte de la trajectoire propre de chaque apprenant au sein d'un système d'activités plus ou moins ouvert/contraint. Si les techniques de fouille de processus ont atteint leur maturité dans le domaine des systèmes d'information où les processus métiers sont bien identifiés, le premier défi de cette thèse consiste à classer les différentes trajectoires que nous aurons identifiées au sein de systèmes d'activité sélectionnés dans un contexte d'apprentissage, afin de pouvoir ensuite déterminer, de manière automatisée, à quelle trajectoire type correspond le parcours et le développement d'un apprenant dans un environnement numérique dépourvu de processus métiers bien identifiés. Il sera ainsi possible d'inscrire son comportement dans la temporalité et la causalité d'une trajectoire, afin de permettre, dans le cadre du développement d'un système intelligent hybride, à l'expert humain de comprendre et visualiser les éléments significatifs de la trajectoire de l'apprenant.

Chaque système de recommandation repose sur un processus en 4 étapes afin de fournir à l'utilisateur actif des éléments intéressants, quelles que soient les techniques d'apprentissage automatique utilisées. Premièrement, il collecte des données d'interaction brutes, également appelées observations. Deuxièmement, il utilise ces données pour déduire une abstraction de haut niveau de l'utilisateur actif dont la représentation est appelée modèle d'utilisateur. Troisièmement, il calcule des recommandations adaptées en fonction des préférences et des attentes de l'utilisateur actif, à partir de ce qu'il a appris dans le modèle utilisateur. Enfin, il doit proposer ces recommandations au bon moment et de la bonne manière via l'interface. Ainsi selon l'algorithme de recommandation, la troisième étape peut nécessiter des entrées supplémentaires, telles qu'une base de connaissances, des modèles de processus ou des représentations du contexte. L'intégration du contexte dans le processus de recommandation comme domaine de recherche émergent appelé CARS, acronyme de « Context Aware Recommender Systems » viendra s'articuler avec les objectifs scientifiques de cette thèse. Dans leur état de l'art, Adomavicius et al. présentent plusieurs approches telles que la modélisation contextuelle, la méthode de pré / post-filtrage pour l'utilisation de facteurs contextuels afin d'adapter la recommandation au contexte de l'utilisateur. Nous nous appuyerons sur ces travaux. De plus, avec la généralisation des usages des smartphones, ainsi que des accès permanents au réseau et ses services, on voit émerger un besoin croissant d'applications capables de services intelligents et personnalisés. Pour ce faire, le système prend en compte un ensemble plus ou moins riche d'informations contextuelles pour offrir une recommandation elle-même contextualisée.

L'objectif de cette thèse est de s'appuyer sur les notions de recommandation de *processus* pour proposer des approches nouvelles de recommandations de parcours complexes implantées dans des compagnons intelligents ou d'applications sur Smartphone. Dans le cadre de cette thèse, nous nous intéresserons aux processus des parcours individuels d'apprentissage et/ou de construction d'un projet personnel ou professionnel. Ceci consistera en l'analyse et la recherche de séquences (sous-processus) d'actions impliquées dans l'achèvement des objectifs des utilisateurs, utilisateurs individuels ou en réseaux. Les résultats s'articuleront alors avec des algorithmes de recommandation qui seront destinés à améliorer l'efficacité du processus de suivi et d'anticipation des actions en lien avec un objectif déclaré ou estimé. Les recherches proposées s'articulent avec les travaux menés dans le cadre du projet KOALA (« Knowledge Aware Learning Assistant ») qu'elles utiliseront

comme terrain d'expérimentation et source des données. KOALA étant une plateforme numérique définissant un espace facilitateur et fédérateur d'engagement installant une dynamique nouvelle de l'occupation de l'espace d'apprentissage et d'enseignement et intégrant la gestion personnalisée ou collaborative des ressources d'apprentissage. Les contributions visées des travaux de thèse permettront de favoriser l'accessibilité, la continuité et la porosité de l'apprentissage en s'appuyant sur la recommandation des ressources adaptées à un projet personnel ou professionnel d'apprentissage réalisés sur un territoire.

Il existe des travaux qui explorent la relation entre le comportement d'apprentissage et les progrès d'apprentissage dans les MOOC, dans le but de mieux comprendre comment les étudiants qui réussissent ou échouent répartissent leurs activités différemment au cours des semaines de cours (van den Beemt et al., 2018). D'autres visent à découvrir les processus d'apprentissage auto-régulés des étudiants lors d'un cours e-Learning en utilisant les techniques de Process Mining (Cerezo et al., 2019). (Romero et al., 2016) ont appliqué le clustering comme tâche de prétraitement pour regrouper les utilisateurs en fonction de leur type d'interactions de cours. Ainsi, ils veulent découvrir les comportements de navigation les plus spécifiques en utilisant uniquement les données en cluster plutôt que l'ensemble de données complet. (Mukala et al., 2015) proposent d'appliquer le Process mining permettant de comprendre les processus d'apprentissage basés sur les traces d'activités des étudiants à partir des journaux des plateformes MOOC. Ils divisent les étudiants en groupes séparés afin d'enrichir leur analyse. Les critères de regroupement sont le type de certificat auquel les étudiants s'inscrivent et le niveau de réussite ou la note finale. Ils ont appliqué l'algorithme fuzzy miner de process mining afin de visualiser et de reproduire le comportement réel des étudiants. (Bakar et al., 2019) veulent identifier les modèles d'apprentissage autorégulé des étudiants avec la technique de fouille de processus en utilisant les données et informations extraites d'un environnement d'apprentissage en ligne.

Les verrous abordés par cette thèse concernent :

- la modélisation et l'analyse d'un parcours individuel personnalisé à partir des traces d'apprentissage ;
- l'analyse et recherche de séquences (sous-processus) d'actions impliquées dans l'achèvement des objectifs des apprenants ;
- le choix et élaboration des méthodes de recommandation adaptées ;
- la validation des différentes méthodes élaborées.

Les contributions attendues des travaux de la thèse concernent à *minima* les points suivants :

- fouille de processus à partir des données fournies depuis une application pédagogique cible ;
- identification des processus nominaux et observés ;
- développement d'algorithmes de recommandation ;
- sélection, test et développement d'algorithmes de recommandation de processus ;
- intégration des algorithmes au sein d'un compagnon intelligent comme avatar de KOALA pour une approche critique et personnalisée de la recherche d'information dans un parcours individuel complexe.

Références :

- SoftLearn: A process mining platform for the discovery of learning paths. Barreiros, Borja Vazquez, Lama Manuel, Mucientes Manuel, Vidal Juan C. Proceedings - IEEE 14th International Conference on Advanced Learning Technologies, ICALT 2014. 373-375.
- Process mining techniques and applications – A systematic mapping study. Garcia Cleiton dos Santos, Meinheim Alex, Faria Junior Elio Ribeiro, Dallagassa Marcelo Rosano, Sato Denise Maria Vecino, Carvalho Deborah Ribeiro, Santos Eduardo Alves Portela, Scalabrin Edson Emilio. Journal Expert Systems with Applications. 260-295.
- Van Harmelen, M. & Workman, D. (2012). CETIS Analytics Series: Analytics for Learning and Teaching.
- Peraya, D. (2019) Entre l'enseignement et la recherche, quelle place pour les Learning Analytics, revue Distances et médiations des savoirs n°27 (DMS27)
- W.M.P. van der Aalst. Business Process Simulation Survival Guide. In J. vom Brocke

- and M. Rosemann, editors, Handbook on Business Process Management 1, International Handbooks on Information Systems, pages 337-370. Springer-Verlag, Berlin, 2015.
- Nowakowski, S., Ognjanović, I., Grandbastien, M., Jovanovic, J., Šendelj, R.: Two Recommending Strategies to Enhance Online Presence in Personal Learning Environments. In: Recommender Systems for Technology Enhanced Learning. Springer, (2014) 227–249.
 - Mohammed Tadlaoui, Système de recommandation de ressources pédagogiques fondé sur les liens sociaux : Formalisation et évaluation, Thèse de Doctorat, co-tutelle Université de Lyon et Université de Tlemcen, juillet 2018.
 - Suire. C., Jean-Caurant. A., Illouz. C., "Ouvrir les Boîtes Noires : un Outil Pédagogique pour une Approche Critique de la Recherche d'Information en Ligne", Digital Humanities 2017 (DH2017), Montréal, 2017, pp. 609-611.
 - <https://www.koala-lms.org/fr/>
 - "Anoptikon", Olivier Auber
 - "Apprendre au XXIème siècle", François Taddei
 - Cohen B, Nowakowski S. « Demain est-il ailleurs ? Odyssée urbaine autour de la transition numérique » FYP Editions 2020
 - "Using online presence data for recommending human resources in the OP4L project" Monique Grandbastien, Suzana Loskovska, Samuel Nowakowski, Jelena Jovanovic, 7th European Conference on Technology Enhanced Learning (EC-TEL 2012), Sep 2012, Sarrebruck, Germany
 - Daouda Sawadogo, Architectures logicielles et mécanismes pour la gestion adaptative et consolidée de ressources numériques dans une application interactive scénarisée, Thèse de Doctorat, La Rochelle Université, juin 2016.
 - Joffrey Leblay, Vers une nouvelle forme d'accompagnement des processus dans les systèmes interactifs, Thèse de Doctorat, La Rochelle Université, août 2019.
 - Adomavicius, G., & Tuzhilin, A. (2011). "Context-aware recommender systems." In Recommender systems handbook (pp. 217-253). Springer, Boston, MA
 - Bakar, M. H. bin A., Ismail, S., & Ali, S. H. S. (2019). *A Process Mining Approach to Understand Self Regulated- Learning In Moodle Environment*. 8(3), 195–200.
 - Baker, R., & Siemens, G. (2014). *Educational data mining and learning analytics*. 253–272. https://doi.org/10.1007/978-1-4614-3305-7_4
 - Bogarín, A., Cerezo, R., & Romero, C. (2018). A survey on educational process mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(1). <https://doi.org/10.1002/widm.1230>
 - Cerezo, R., Bogarín, A., Esteban, M., & Romero, C. (2019). Process mining for self-regulated learning assessment in e-learning. *Journal of Computing in Higher Education*. <https://doi.org/10.1007/s12528-019-09225-y>
 - Castagnos, S., L'Huillier, A., & Boyer, A. (2015, November). "Toward a robust diversity-based model to detect changes of context." In 2015 IEEE 27th International Conference on Tools with Artificial Intelligence (ICTAI) (pp. 534-541). IEEE
 - Labarthe, H., & Luengo, V. (2016). *L'analytique des apprentissages numériques*.
 - Long, P., & Siemens, G. (2011). Penetrating the fog: Analytics in learning and education. *Educause Review*. <https://doi.org/10.1177/0894318416647779>
 - Mukala, P., Buijs, J., Leemans, M., & van der Aalst, W. (2015). Exploring Students' Learning Behaviour in MOOCs using Process Mining Techniques. *Computing Conference*, 1–12. <http://bpmcenter.org/wp-content/uploads/reports/2015/BPM-15-10.pdf>
 - Romero, C., Cerezo, R., Bogarín, A., & Sánchez-Santillán, M. (2016). EDUCATIONAL PROCESS MINING: A TUTORIAL AND CASE STUDY USING MOODLE DATA SETS. In *Data Mining And Learning Analytics: Applications in Educational Research* (pp. 3–28). <https://doi.org/10.1002/9781118998205>
 - Trabelsi, M., Suire, C., Morcos, J., & Champagnat, R. (2019). *Fouille de processus auto-définis : cas d'étude d'un moteur de recherche d'une bibliothèque numérique*. August. <https://doi.org/10.3166/INFORSID..1-16>
 - van den Beemt, A., Buijs, J., & van der Aalst, W. (2018). Analysing structured learning behaviour in Massive Open Online Courses (MOOCs): An approach based on process

mining and clustering. *International Review of Research in Open and Distance Learning*, 19(5), 38–60. <https://doi.org/10.19173/irrodl.v19i5.3748>

Travail demandé au doctorant

Préciser les tâches qui seront confiées au doctorant (programme de travail)

Les travaux s'appuieront sur les domaines du *process mining*, qui regroupe des techniques de gestion de processus métier basées sur la journalisation des événements et l'analyse des traces d'activité et les *systèmes de recommandation*, qui représentent une famille d'outils et de méthodes permettant, à partir de diverses données collectées dans l'environnement de l'utilisateur, de proposer une ou plusieurs suggestions compte tenu de ces données.

Le thésard devra, plus précisément :

- étudier l'application des algorithmes de la fouille de processus aux Learning Analytics afin d'extraire des modèles de trajectoires ;
- définir une méthode de classification des trajectoires ;
- choisir et élaborer des méthodes de recommandation adaptées en utilisant des techniques d'intelligence artificielle ;
- valider les méthodes choisies sur le cas d'études développé au laboratoire.

Profil recherché :

Formation générale en informatique de niveau M2 Recherche avec des connaissances en génie logiciel et en modélisation des processus métiers. Des notions en formalisation mathématique et en fouille de données seraient un plus.

Un bon niveau d'anglais est requis.

Modalités de candidature :

- remplir et déposer le dossier de candidature avec les pièces mentionnées dans le formulaire de candidature.
- contacter par mail un des 2 co-encadrants de La Rochelle Université (coordonnées ci-dessus) pour signaler le dépôt de candidature.