Projet HUBBLE

HUman oBservatory Based on anaLysis of e-LEarning traces.

Projet ANR-14-CE24-0015
Défi Principal : Société de l'information et de la communication. Axes : « Formation et éducation » et « Interactions des mondes physiques, de l'humain et du monde numérique »
Défi secondaire : Sociétés innovantes, intégrantes et adaptatives. Axe : « formation et éducation»

Porteur scientifique :
Vanda Luengo
Univ. Grenoble Alpes, LIG, F-38000 Grenoble, France CNRS, LIG, F-38000 Grenoble, France Inria1

Partenaires

Equipe MeTAH, Laboratoire LIG, Université Grenoble Alpes.
Laboratoire LINA, Université de Nantes,
Equipe Silex, Université Claude Bernard, Lyon 1.
Equipe IEIAH, Université du Mans (MAINE)
Entreprise OpenClassrooms
Equipe EduTICE, Ecole normale supérieure de Lyon
Laboratoire STEF, Ecole normale supérieure de Cachan
Laboratoire LabSTICC, Institut de Mines Télécom, Télécom Bretagne.

Résumé

Le projet HUBBLE propose la création d'un observatoire pour la construction et le partage de processus d'analyse des traces e-learning massives. HUBBLE permettra aux différents acteurs (enseignants, apprenants, concepteurs, administrateurs ou politiques) d'analyser et d'expliquer les phénomènes d'enseignement et d'apprentissage avec des environnements de type e-learning. Les processus d'analyse construits accompagneront la prise de décisions des acteurs intervenant dans le système d'enseignement et d'apprentissage, ils devront être utilisables et redéployées dans leurs environnements. Ils guideront, par ailleurs, le chercheur en e-learning pour la production de concepts, de modèles et d'indicateurs.
HUBBLE prendra appui sur les plateformes existantes permettant la construction, le partage et la gestion de processus d'analyse à partir des traces disponibles dans l'observatoire. HUBBLE considérera les problèmes éthiques concernant les traces et les analyses résultantes des processus.
Le projet HUBBLE a vocation à fédérer des équipes de recherche en informatique et en sciences humaines, avec une volonté affichée de développer un observatoire national. Il est organisé autour des cas d'études portées par les partenaires du projet. Ces cas d'études permettront de concevoir et développer les processus d'analyse qui seront ensuite partagées.
Au-delà de ses résultats expérimentaux ou technologiques, le projet vise à favoriser la construction d'une communauté nationale sur la thématique des "Learning Analytics" afin de pouvoir se positionner au niveau européen et international.

Contexte, positionnement et objectif de la proposition détaillée

Le projet HUBBLE propose la création d'un observatoire pour la construction et le partage de traces massives d'e-learning, de leurs processus d'analyse et de leurs contextes d'usage. Il s'agit d'un projet

fédérateur et structurant autour des environnements d'e-learning massif (dont les MOOCS, mais aussi les environnements de type serious games ou des simulations) comportant une dimension évaluative longitudinale pour comprendre les processus éducatifs afférents. Le projet HUBBLE a vocation à fédérer des équipes de recherche en informatique et en sciences humaines, avec une volonté affichée de développer un observatoire national.
Des processus d'analyse seront construits pour assister la prise de décision des acteurs intervenant dans le système d'enseignement et d'apprentissage (acteur décisionnel). D'autres processus seront construits pour assister le chercheur en e-learning à la production de concepts, modèles et indicateurs permettant d'analyser et d'expliquer les phénomènes d'enseignement et/ou apprentissage avec des environnements e-learning. Les processus construits pour assister la prise de décision des acteurs décisionnels devront être redéployés dans les environnements habituels de ces acteurs.
Nous appelons acteur décisionnel dans le domaine du e-learning toute personne utilisant de façon directe ou indirecte des environnements d'e-learning. Ainsi, le concepteur d'un MOOC, l'apprenant utilisant le MOOC, un enseignant utilisant un système e-learning, le concepteur du système lui-même ou bien son administrateur sont des acteurs décisionnels. Le responsable pédagogique d'une formation, ou encore les responsables administratifs et politiques d'une institution d'enseignement utilisant des systèmes d'e-learning sont également des acteurs décisionnels.
Nous appelons processus d'analyse l'application d'une succession d'outils supportant le traitement et/ou la visualisation interactive des traces. Il pourra mobiliser plusieurs outils de collecte, d'analyse ou de visualisation. Un processus d'analyse doit être caractérisé par rapport à son contexte d'analyse et son contexte d'usage.
L'observatoire prendra appui sur des plateformes existantes permettant la collecte, le partage et la réutilisation des traces d'e-learning ainsi que des processus d'analyse de ces traces. L'observatoire s'intéressera aux problèmes éthiques concernant le partage des traces et des analyses résultantes, en créant un comité d'éthique pour ce projet.

2.1 Contexte

Rendre compte en temps réel ou en temps différé des usages de dispositifs de e-learning représente un enjeu majeur tant du point de vue de la recherche que du point de vue opérationnel. Ainsi, un décideur souhaitant mesurer l'impact et l'audience d'une formation, ou une équipe pédagogique souhaitant modifier une formation, pourront être intéressés par une typologie des différents types d'apprenants : auditeurs libres consultant de manière passive les ressources mises à disposition, participants actifs impliqués dans les différentes activités proposées, ou décrocheurs se désengageant de la formation. Autre exemple, pour piloter une formation ou personnaliser un parcours d'apprentissage, un praticien pourra souhaiter disposer d'indicateurs permettant de guider des interventions globales ou ciblées, auprès de décrocheurs par exemple. Un apprenant pourra être intéressé par la visualisation de sa trajectoire individuelle au sein de la formation, et son positionnement par rapport aux autres apprenants (*awareness*).
Enfin, un chercheur pourra exploiter les traces d'utilisation laissées par les apprenant et enseignant afin de produire des nouvelles connaissances, modèles et outils pour comprendre et accompagner les phénomènes d'enseignement et apprentissages avec l'e-learning. Par exemple un chercheur peut s'intéresser à l'activité (à travers les traces et les observations) des enseignants, lorsqu'ils conçoivent les exercices via des outils auteurs, afin d'acquérir des connaissances permettant de proposer des rétroactions pertinentes ou encore d'identifier des connaissances du domaine enseigné ou de diagnostic permettant d'assister l'enseignant lors de sa création d'exercices.
Dans ce contexte, le projet HUBBLE propose la création d'un observatoire pour la construction et le partage de traces massives d'e-learning, de leurs processus d'analyse et de leurs contextes d'usage.
Le projet HUBBLE s'intéresse ainsi aux enjeux technologiques et méthodologiques de l'axe "formation et éducation" (des défis *Société de l'information et de la communication* et du défi *Sociétés innovantes, intégratives et adaptatives*). Les objectifs technologiques du projet sont centrés sur des services collaboratifs permettant de capitaliser l'intelligence collective tel qu'indiqué dans l'axe "interactions des

mondes physiques, de l'humain et du monde numérique" (axe 3.7.2.4 de l'appel). L'enjeu d'évaluation et d'analyse lié aux MOOCs, et de façon plus générale aux environnements d'e-learning, est également central dans le projet HUBBLE (liée à l'axe 3.8.3 du défi Sociétés innovantes, intégratives et adaptatives).
En ce qui concerne les enjeux méthodologiques et technologiques, le projet HUBBLE cherche à partager les processus d'analyse de façon à :
1) Permettre d'expérimenter de nouvelles pédagogies (Inquiry-based, connectivism, serious games, collaborative, etc.) et de nouvelles formes d'interaction (personnalisation et adaptation des ressources - contenus, outils, personnes ou groupes, etc. - aux utilisateurs).
2) Permettre de tester des hypothèses de recherche à partir des scénarios « d'analyse chercheur ».
3) Permettre l'évaluation longitudinale des nouveaux dispositifs (comme par exemple les MOOCS) à partir de "l'histoire" et de l'évolution temporelle des traces, des processus d'analyses et de leurs effets sur l'ingénierie pédagogique.
4) Permettre la reproductibilité des analyses sur plusieurs corpus de traces afin d'améliorer leur qualité et leur pertinence ou leur domaine de validité.
5) Permettre des analyses d'un même corpus selon plusieurs points de vue.
6) Autoriser des modifications plus ou moins importantes des processus d'analyse pour étudier la solidité des résultats (seuils, clustering, catégorisations, etc.).
En ce qui concerne l'analyse et l'évaluation, les formations de type e-learning introduisent de nouvelles formes d'enseignement et d'apprentissage (tel que le dispositif de pédagogie inversée en médecine à Grenoble) ou des nouveaux dispositifs de type e-learning (tel que les MOOCS). Il est donc nécessaire de :
1) Comprendre les nouveaux phénomènes émergents.
2) Etudier les pédagogies innovantes dans des contextes massifs (connectivisme, démarche par investigation, etc.), utilisant des technologies interactives.
3) Rendre compte aux différents acteurs décisionnels (enseignants, formateurs, tuteurs…) de l'évolution de la situation d'apprentissage (indicateurs, tableaux de bord, retours d'expérience, etc.) et les assister dans leurs décisions.

2.2 Etat de l'art et positionnement

Il existe un corpus important sur l'analyse de fichiers log et de traces [Dwyer et al. 1999, LeBlanc et al. 1990, Lee et al. 1999, Mansouri-Samani & Sloman 1997, Zhang & Wenke 2000]. Historiquement, l'analyse de logs s'est d'abord focalisée sur le fonctionnement des programmes, à des fins de surveillance et de débogage. Mais progressivement, leur potentiel pour analyser l'activité des utilisateurs a été reconnue. Les techniques de fouille de données et d'apprentissage automatique ont alors été utilisées pour la découverte de processus dans les activités médiées par l'informatique [Cook & Wolf 1998, Song et al. 2009], et plus récemment pour l'identification de communautés dans les réseaux sociaux basée sur des motifs d'interaction entre utilisateurs [Sachan et al. 2012].
L'analyse statistique et/ou synthétique n'est pas la seule exploitation possible des traces. Les traces d'activité peuvent aussi être considérées comme un réservoir d'expériences individuelles réutilisables dans des contextes similaires, soit à l'identique soit après adaptation [Cordier et al. 2013]. Cette hypothèse sous-tend un certain nombre de travaux, visant pour certains à fournir à l'utilisateur des recommandations basées sur des expériences passées [Groza et al. 2007, Hug et al. 2012], pour d'autres à suivre la progression des apprenants dans des applications de e-learning [Poltrack et al. 2012]. Avec le développement des dispositifs mobiles et des capteurs embarqués, les pratiques de traçage de l'activité quotidienne voient le jour : lifelogging [O'Hara et al. 2008] et *quantified self* [Wolf et al. 2010] visent une meilleure connaissance de soi et une remémoration facilitée des événements passés.
Selon leurs objectifs, les différentes approches citées ci-dessus requièrent l'enregistrement de différents événements dans les traces qu'ils utilisent respectivement. À l'exception du *quantified self* (qui met l'accent sur des données numériques provenant de capteurs), la plupart de ces travaux s'appuient sur des événements de haut niveau (par exemple : lancer une application, ouvrir un fichier). Il en résulte que les outils

disponibles pour la collecte de traces sont le plus souvent limités aux événements de haut niveau[1], ou au contraire à des événements de très bas niveau (frappe clavier, mouvements de souris) dépourvus de contexte[2]. Dans ce projet, nous opterons donc pour une approche plus riche permettant de mettre en relation ces différents niveaux [Ginon et al. 2013]. Pour les applications Web, une difficulté supplémentaire est posée par l'éclatement des informations entre le client et le serveur, ce qui nécessite une collecte distribuée [May et al. 2011, Le et al. 2013].

Dans le cadre d'une démarche par investigation ou connectiviste [Siemens 2005] pour des MOOCs, ils se basent en général sur un site d'accueil simple pour permettre l'inscription, l'affichage des thématiques proposées, les premiers échanges (notamment pour se présenter) au travers d'un forum, (qui peut être aussi une plateforme MOOC classique, type Open Edx, coursera, Openclassroom, etc.). Néanmoins, l'essentiel des échanges entre participants se développe de manière distribuée sur les outils du web2 (blogs, réseaux sociaux, outils d'annotation, wikis, outils de production collaborative, etc.). L'apprenant se constitue alors un environnement d'apprentissage personnel de l'élève (PLE) [CHATTI et al. 07], point d'entrée de son environnement social d'apprentissage. L'outil essentiel d'un tel MOOC est alors de pouvoir agréger les traces d'activités sous forme de flux. Si des agrégateurs de flux peuvent suffire, des outils complémentaires comme EduFeedr [POLDOJA et al. 10] ou gRSShopper [DOWNES 10] ont été proposés. Ce dernier permet de configurer des vues personnalisées, selon une perspective de PLE. Ces outils permettent en effet de rendre visible et de partager l'activité générée par le cours, ce qui est indispensable à la dynamique du cours et à la sensation d'appartenance à une communauté d'apprentissage. Ces derniers sont néanmoins insuffisants pour filtrer de manière efficace et pertinente les flux important d'informations et de traces d'activités pour chaque type d'utilisateurs.

Par ailleurs, il est bien sûr toujours possible, comme c'est encore souvent le cas en sciences humaines, de procéder à une collecte de trace manuelle (prise de notes) ou semi-automatisée (captation audio ou vidéo). Bien que ce mode de collecte permette de palier certaines difficultés techniques, elle est coûteuse en moyens humains. Par ailleurs, les traces produites sont difficilement interprétables par la machine, même si des outils d'annotation vidéo, par exemple, visent à combler ce manque [Aubert et al. 2012].

Dans le projet HUBBLE, l'objectif est d'analyser des traces formalisées collectées automatiquement, en gardant la possibilité de les enrichir avec des traces manuelles et/ou des traces issues de captation audio ou vidéo. Par ailleurs il est possible d'augmenter les traces, en rajoutant des observables non collectés dans les plateformes e-learning afin d'observer certains phénomènes d'interaction liées à l'enseignement ou à l'apprentissage [Sharma & al 2013].

Dans le domaine du e-learning, la collecte et l'analyse de traces sont des domaines émergeants représentés par la communauté d'Educational Data Mnining (EDM - http://www.educationaldatamining.org/) et de learning analytic (LAK - http://www.solaresearch.org/). Plusieurs travaux s'intéressent à la capitalisation de traces. DATAShop [KOEDINGER & al 2010] est l'une des plateformes les plus connues et utilisée dans EDM. Elle met à disposition des traces avec l'objectif spécifique de trouver et tester des modèles d'apprenant. Les traces sont des traces d'interaction avec des tuteurs intelligents de type cognitive tutor .

En France le projet MULCE [REFFAY & al. 2012] est représentatif de l'effort autour de la collecte et du partage de traces. Ce projet propose la diffusion de corpus d'apprentissage, appelés LETEC, avec comme objectif la mise à disposition non seulement des données résultat d'une formation, mais également des données du contexte, c'est-à-dire caractérisant le dispositif de formation ainsi que le dispositif de recherche. L'objectif est de rendre ce type de données visibles, partageables et réutilisables.

Ce deux projets ne capitalisent pas les processus d'analyse mais proposent des liens vers des outils libres d'analyse développées par des chercheurs. Les auteurs de ces projets soulignent l'importance de réaliser des connections entre leur plateforme et ce type d'outils. Ainsi, il existe des plateformes de traitement de traces génériques (comme Weka pour le data mining ou R pour la statistique), mais certains auteurs [ROMERO & al. 2007] pointent la nécessité de concevoir des processus d'analyse propres au domaine d'e-learning.

[1] Par exemple http://dev.nepomuk.semanticdesktop.org/ ou http://intersectalliance.com/snareagents/

[2] Par exemple https://github.com/gurgeh/selfspy ou http://www.mykeylogger.com/

De ce point de vue (outils d'analyse), la différence principale des deux communautés (EDM et LAK) sont les cibles des productions. En effet les algorithmes et outils produits dans EDM sont principalement utilisés par les systèmes informatiques alors que la communauté LAK s'intéresse principalement à fournir des outils pour les utilisateurs humains du domaine. De plus, EDM met l'accent sur la découverte automatique des connaissances alors que LAK met l'accent sur la découverte à partir du jugement humain [Siemmens & Baker 2012].
Des algorithmes produits dans la communauté d'EDM pourraient être de grande utilité s'ils sont pris en compte dans le processus d'analyse comme des opérateurs. Autrement dit, ces algorithmes peuvent être associés à d'autres opérateurs au sein d'un processus. Par exemple, les algorithmes qui détectent un type de comportement de l'apprenant [Baker & al. 2013], ou qui diagnostiquent les connaissances des apprenants [Desmarais & al. 2012], peuvent être associés à des visualisations dans un processus d'analyse, pour pouvoir informer les acteurs décisionnels. En ce qui concerne le domaine de LAK, ils se basent principalement sur des outils analytic du commerce mais qui ne tiennent pas compte des particularités du domaine [Siemmens & Baker 2012], comme par exemple l'analyse à plusieurs échelles (apprenant ou classe, leçon ou programme). Le domaine s'est ainsi investi très récemment dans la production de type d'outils adaptés au domaine du e-learning, utiles et utilisables pour les acteurs du e-learning.
Dans le cadre du projet HUBBLE, nous souhaitons associer ces deux approches. Le consortium du projet associe ainsi des équipes qui sont parmi les acteurs principaux dans le domaine des EIAH et qui s'intéressent à l'usage des plateformes e-learning ainsi qu'à la collecte et l'analyse des traces d'e-learning. Le consortium est de grande taille car il implique des partenaires compétents dans tous les éléments de la chaine pour la production d'analyses pertinentes dans le domaine. Cette chaine va de la recherche avec les acteurs décisionnels de terrain (avec des apprenants, enseignants, formateurs, plateformes) jusqu'à la proposition des outils de collecte, de traitement et de visualisation, en passant par la proposition des modèles d'apprenant et de rétroaction ou des modèles sémantiques. Le projet est donc compatible avec des démarches de type design-based research (Design-Based Research Collective, 2003 ; Wang & Hannafin, 2005) qui articulent des phases de co-conception avec les acteurs de terrain et de collecte de données permettant de nourrir les travaux de recherche. Ainsi, tel que montré dans la description du consortium, certains partenaires sont spécialistes dans le travail de terrain ([Bruillard 2012], [Sanchez 2013],[Emin et al 2014]), ou dans les plateformes de collecte ([Reffay et al. 2012], [Champin & al. 2013], [Iksal 2011]) et d'analyse ([Bouhineau et al. 2013], [May et al. 2011]), ou dans l'analyse des données pour la production des modèles d'apprenant ([Lallé & al. 2013], [Luengo 2011]) ou dans la proposition des modèles sémantiques pour la gestions de traces dans le domaine du e-learning ([Cheniti l. Garlatti S. 2013], [Gilliot 2013],[Setouti & al. 2011]). L'enjeu ici est donc de mutualiser les différentes compétences, méthodes et outils pour permettre la capitalisation et réutilisation des processus d'analyse et de promouvoir ce type de recherche en France, actuellement très dynamique au niveau l'international.

Le projet est *ambitieux* à plusieurs titres :

- Construction et diffusion de processus d'analyses. Dans ce cadre, il importe que certains processus puissent devenir utilisables pour des acteurs décisionnels. Ce passage n'est pas automatique, nous travaillerons dans la construction de cette « chaine éditoriale » de processus.
- Recherche pluridisciplinaire émanant de différentes disciplines associées aux environnements informatiques pour l'apprentissage humain.
- Gestion de données respectueuse de la vie privée. Il sera nécessaire d'établir des méthodes, par le biais d'un comité d'éthique, garantissant que toute trace et toute analyse de traces respecte les règles d'éthique établies par le comité.

Le projet est *novateur* dans plusieurs dimensions :

- Partage de processus d'analyse. Le projet propose le partage, non seulement des traces comme dans la plupart des plateformes, mais aussi des processus d'analyse de ces traces.
- Intégration dans la complexité du Web. Du point de vue informatique, le projet doit tenir compte de l'aspect massif, « cloud computing » et sémantique des traces et des processus pour proposer des solutions utilisables par tous les acteurs concernés.
- Analyse en temps réel des traces d'activité. Les processus d'analyse doivent permettre l'analyse

non seulement de traces fermées, i.e. traces qui ont été collectés dans le passé, mais également des traces ouvertes, i.e. traces qui arrivent au fur et à mesure de l'analyse.

- Une expérience utilisateur étudiée pour chaque acteur. Du point de vue de l'expérience utilisateur, le projet doit tenir compte de l'utilisabilité et de l'utilité des processus d'analyse pour chacun des acteurs décisionnel du e-learning.

Le projet HUBBLE s'inscrit ainsi dans l'agenda stratégique présenté en octobre 2013 par la Ministre de l'Enseignement supérieur et de la Recherche Geneviève Fioraso. Parmi les actions mises en avant, nous pouvons relever l'action n°11 consistant à donner une impulsion forte à la recherche sur la pédagogie numérique et notamment à la recherche dans l'e-éducation, l'enjeu étant d'améliorer la structuration d'une communauté de recherche pluridisciplinaire dans le domaine de l'e-éducation.
Par ailleurs, au-delà de ses résultats expérimentaux ou technologiques, le projet vise à favoriser la construction d'une communauté nationale sur la thématique des "Learning Analytics", dont l'importance va en grandissant, afin de pouvoir se positionner au niveau européen et international (c.f. lot 7 du projet).

2.3 Objectifs

Les objectifs scientifiques du projet sont :

1) La proposition de modèles, langages et méthodes pour assister des utilisateurs non informaticiens dans l'interprétation de traces massives issues d'une activité de e-learning. Ceci implique :
   a) Des moyens de collecte et de modélisation de traces. Par exemple : répondre à la question de comment permettre l'importation de traces existantes dans les plateformes d'analyse du point de vue technique et éthique.
   b) Des moyens d'analyse permettant de décrire et mettre en œuvre les transformations/requêtes/calculs d'indicateurs/visualisation sur les traces.
   c) Des moyens pour décrire les processus d'analyse pour différents acteurs décisionnels. Par exemple : à partir des processus d'analyse, répondre à la question de comment créer des tableaux de bord adaptés à plusieurs décideurs (créateur de ressources, apprenant, tuteur, responsable politique, chercheurs, etc.).
2) La construction de nouvelles connaissances :
   a) Sur des phénomènes liés à l'usage des systèmes d'e-learning ou des nouvelles pédagogies associées. Comme par exemple : la nature de l'engagement dans les MOOCS ; la nature du tutorat dans des environnements massifs ; les méthodes d'évaluation (auto-évaluation, évaluation par les pairs, évaluations semi-automatiques, etc.) ;
   b) Sur les modèles, méthodes, stratégies dans le domaine, en particulier : les modèles de l'apprenant, de rétroaction, des groupes sociaux d'interaction, du contexte et les méthodes d'adaptation/personnalisation.

Les objectifs technologiques du projet sont :

1) Mise en commun d'expériences sur les plateformes d'observation et d'analyse afin de permettre le partage, la capitalisation et l'usage de :
   a) traces issues d'environnements e-learning avec leur contexte d'usage (environnement utilisé, pédagogie, type d'acteurs, etc.) pour assurer une certaine réutilisation.
   b) outils permettant d'analyser et de visualiser ces traces.
   c) processus d'analyse avec leur contexte et leurs résultats tels que des indicateurs ou des modèles.
2) L'utilisabilité des processus d'analyse pour des utilisateurs qui souhaitent les réutiliser.
3) Le redéploiement de processus d'analyse pour la prise de décision des acteurs institutionnels dans leurs environnements habituels.
4) Le respect des règles éthiques définies par le comité d'éthique qui sera créé lors du projet.

Références

Aubert, O., Prié, Y., & Schmitt, D. Advene, a versatile hypervideo authoring tool. ACM Symposium on Document Engineering, Paris. (2012).

Baker, R.S.J.d., Corbett, A.T., Roll, I., Koedinger, K.R., Aleven, V., Cocea, M., Hershkovitz, A., de Carvalho, A.M.J.B., Mitrovic, A., Mathews, M. (2013) Modeling and Studying Gaming the System with Educational Data Mining. In Azevedo, R., & Aleven, V. (Eds.) *International Handbook of Metacognition and Learning Technologies*. pp. 97-116. New York, NY: Springer.

Bouhineau, D., S. Lalle, V. Luengo, N. Mandran, M. Ortega, and C. Wajeman. Share data treatment and analysis processes inTechnology enhanced learning. In Workshop Data Analysis and Interpretation for Learning Environments, Autrans, France, 2013. Alpine Rendez-Vous 2013.

Bruillard Éric (2012). Le déploiement des ENT dans l'enseignement secondaire : entre acteurs multiples, dénis et illusions, Revue Française de Pédagogie, 177, oct.-déc. 2011, p. 101-130.

P-A. Champin, A. Mille, Y. Prié. Vers des traces numériques comme objets informatiques de premier niveau : une approche par les traces modélisées. Intellectica (59):171-204. 2013.

Cheniti L., GARLATTI Serge. (2013) "Semantic Web Technologies for Supporting Pervasive Peer Assessment", TEFA 2014 : Workshop on Technology-Enhanced Formative Assessment (TEFA) 2013,http://www.kbs.uni-hannover.de/tefa2013.html : 17-18 septembre 2013, Paphos, Cyprus, Chypre, 2013

Cook, J. E., Wolf, A. L. :. Discovering Models of Software Processes from Event-Based Data. In: ACM Transactions on Software Engineering and Methodology, 7(3), 215–249. (1998)

Desmarais, Michel C., and Ryan S. J. d Baker. 2012. 'A review of recent advances in learner and skill modeling in intelligent learning environments.' User Modeling and User-Adapted Interaction 22(1-2), pp. 9–3

Design-Based Research Collective. 2003. 'Design-based research: An emerging paradigm for educational inquiry'. Educational Researcher,, 32(1), pp 5-8.

Dwyer, M. B., Avrunin, G. S., & Corbett, J. C. Patterns in property specifications for finite-state verification. In: Software Engineering, 1999. Proceedings of the 1999 International Conference on (pp. 411-420). IEEE. (1999).

Emin-Martínez, C. Hansen, M. J. Rodríguez-Triana, B. Wasson, Y. Mor, M. Dascalu, R. Ferguson, J.-P. Pernin, Towards Teacher-led Design Inquiry of Learning, Revue eLearning Papers, 01/2014

GILLIOT Jean-Marie, GARLATTI Serge, REBAÏ Issam, BELEN-SAPIA Maria Le concept de iMOOC pour une ouverture maîtrisée. EIAH 2013 : atelier thématique MOOC - Massive Open Online Course - État des lieux de la recherche francophone, 28-28 mai 2013, Toulouse, France, 2013.

Ginon, B., Champin, P-A. , & Jean-Daubias, S. Collecting fine-grained use traces in any application without modifying it. Workshop EXPPORT from the conference ICCBR, New York, USA. (2013)

Groza,, T.; Handschuh, S., Möller, K., Grimnes, G., Sauermann, L., Minack, E., Mesnage, C., Jazayeri, M., Reif, G., Gudjonsdottir. R. The NEPOMUK Project -- On the way to the Social Semantic Desktop. In: Proceedings of the Third International Conference on Semantic Technologies (I-SEMANTICS 2007), Graz, Austria. (2007)

Hug, C., Deneckere, R., Salinesi, C.: Map-TBS: Map process enactment traces and analysis, In: International Conference on Research Challenges in Information Science (RCIS), Valencia : Espagne (2012).

Iksal S. Tracks Analysis in Learning Systems: A Prescriptive Approach. In International Journal for e-Learning Security (IJeLS), Volume 1, Issue 1, 2011

Koedinger, K.R., Baker, R.S.J.d., Cunningham, K., Skogsholm, A., Leber, B., Stamper, J.. A Data Repository for the EDM community: In *Handbook of Educational Data Mining*. Boca Raton, FL: CRC Press. 2010.

S. Lalle, J. Mostow, V. Luengo, and N. Guin. Comparing Student Models in Different Formalisms by Predicting their Impact on Help Success. In AIED 2013: 16th International Conference on Artificial Intelligence in Education. IOS press, Juillet 2013

Le, A.H. , Lefevre, M. , Cordier, A. , Skaf-Molli, H. Collecting Interaction Traces in Distributed Semantic Wikis. 3rd International Conference on Web Intelligence, Mining and Semantics (WIMS '13), Madrid, Spain. pp. 21:1-21:11. ACM New York, NY, USA. ISBN 978-1-4503-1850-1. (2013)

LeBlanc, T. J., Mellor-Crummey, J. M., & Fowler, R. J. Analyzing parallel program executions using multiple views. Journal of Parallel and Distributed Computing, 9(2), 203-217. (1990)

Lee, W., Stolfo, S. J., & Mok, K. W. A data mining framework for building intrusion detection models. In Security and Privacy, 1999. Proceedings of the 1999 IEEE Symposium on (pp. 120-132). IEEE. (1999)

May M., George S., Prévôt P., TrAVis to Enhance Students' Self-Monitoring in Online Learning Supported by Computer-Mediated Communication Tools, *International Journal of Computer Information Systems and Industrial Management Applications (IJCISIM)*, Vol. 3/4, 2011, p. 623–634.

Mansouri-Samani, M., Sloman, M.: GEM: a generalized event monitoring language for distributed systems. In: Distruted Systel Engineering, 4(2). (1997)

Zhang, Y., and Wenke L. Intrusion detection in wireless ad-hoc networks. In Proceedings of the 6th annual international conference on Mobile computing and networking, pp. 275-283. ACM. (2000)/

Michelet, S.. and Luengo, V.. Prise en compte des contradictions intra-apprenant dans le diagnostic étude de cas : DiagElec un diagnostic informatique. Sciences et Technologies de l'Information et de la Communication pour l'Éducation et la Formation, 19:407--436,

O'Hara, K., Tuffield, M. M., & Shadbolt, N. Lifelogging: Privacy and empowerment with memories for life. Identity in the Information Society, 1(1), 155-172. (2008)

Poltrack, J., Hruska, N., Johnson, A., & Haag, J. The Next Generation of SCORM: Innovation for the Global Force. In The Interservice/Industry Training, Simulation & Education Conference (I/ITSEC) (Vol. 2012, No. 1). National Training Systems Association. (2012)

Reffay C., Betbeder M. and Chanier T. Multimodal learning and teaching corpora exchange: lessons learned in five years by the Mulce project. In *Int. Journal Technology Enhanced Learning*, Vol. 4, Nos. 1/2, pp.11–30. 2012

Romero C. and Ventura S. Educational Data Mining: « A Survey from 1995 to 2005». *Expert Systems with Applications*, 33(1), 2007, pp. 135-146.

Sachan, M., Contractor, D., Faruquie, T. a., Subramaniam, L. V.: Using content and interactions for discovering communities in social networks. In: Proceedings of the 21st international conference on World Wide Web - WWW '12, pp 331-341. (2012).

Sanchez, E., Ney, M., & Labat, J. (2011). "Jeux sérieux et pédagogie universitaire : de la conception à l'évaluation des apprentissages". Revue internationale des technologies en pédagogie universitaire, 8(1-2), 46-57.

Settouti L., Guin N. , Luengo V., and Mille A. Adaptable and Reusable Query Patterns for Trace-Based Learner Modelling. In EC-TEL, pp 384-397, 2011.

Sharma k., Jermann P., M.-A. Nüssli and P. Dillenbourg. Understanding Collaborative Program Comprehension: Interlacing Gaze and Dialogues. Computer Supported Collaborative Learning (CSCL 2013), Madison, Wisconsin, USA, 2013.

Siemens, G. (2005). Connectivism: a learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning*, 2, 3-10.

Siemens, G., Baker, R.S.J.d. (2012) Learning Analytics and Educational Data Mining: Towards Communication and Collaboration. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*.

Song, M., Günther, C. W., Van der Aalst, W.: Trace Clustering in Process Mining. In M. Van der Aalst, M. et al. (Eds.), Business Process Management Workshop (pp. 109–120). Springer Berlin Heidelberg. (2009)

Wang, F, & Hannafin, M. J. 2005. Design-based research and technology-enhanced learning environments. Educational Technology Research and Development. Educational Technology Research and Development, 53(4), pp 5-23.

Wolf, G., Carmichael, A., & Kelly, K. The quantified self. TED. (2010) http://www.ted.com/talks/gary_wolf_the_quantified_self.html