



Les mots de la recherche sur les EIAH, enjeux et questions

► **Nicolas BALACHEFF** (LIG, Université Grenoble Alpes, CNRS, Grenoble INP)

■ **RÉSUMÉ** • Le contexte international et la nature essentiellement pluridisciplinaire de la recherche sur les Environnements informatiques pour l'apprentissage humain (EIAH) soulèvent des problèmes d'intercompréhension entre les disciplines, et entre les différentes langues à travers le filtre d'une langue de travail commune - l'anglais. Au-delà des problèmes de mots, c'est la compréhension partagée des concepts qui est en question. Après une présentation des principales caractéristiques de ces problèmes, cet article présente un projet de dictionnaire des termes et expressions de la recherche sur les EIAH et en décrit l'état actuel de réalisation.

■ **MOTS-CLÉS** • dictionnaire de la recherche sur les EIAH, vocabulaire de la recherche sur les EIAH, pluridisciplinarité

■ **ABSTRACT** • *The international context and the fundamentally multidisciplinary nature of Technology Enhanced Learning (TEL) research raise problems of inter-comprehension between disciplines, between different languages through the filter of a common working language - English. Beyond problems with the translation of words, what is in question is a shared understanding of concepts. After a presentation of the main features of these problems, this paper presents the project of a dictionary of TEL research terms and expressions and describes its current state of realization.*

■ **KEYWORDS** • *TEL research dictionary, TEL research vocabulary, multidisciplinary*

1. Le dictionnaire, instrument pour la construction d'une référence commune

La recherche sur les environnements informatiques pour l'apprentissage humain (EIAH) est, comme toutes les activités scientifiques contemporaines, largement internationale. La langue dominante de communication est l'anglais. Elle est aussi la langue de travail des projets internationaux, notamment dans la communauté européenne, alors que les recherches sont pensées et conduites dans les langues nationales des différents pays. Par ailleurs, la recherche sur les EIAH est par nature pluridisciplinaire. Elle ne peut progresser que par l'intégration profonde de plusieurs disciplines que l'on peut, dans une forme de raccourci, identifier à l'informatique et aux sciences humaines. Au sein de chacune, il s'agit de plusieurs spécialités qui elles-mêmes diffèrent par leurs concepts et leurs méthodes. Ces deux caractéristiques, internationalisation et pluridisciplinarité, sont à la source de ce qui est le plus souvent repéré comme un problème de communication, mais qui est en fait bien plus critique que cela. Les difficultés rencontrées lors des échanges (collaborations, exposés, articles) ne sont pas que des « problèmes de mots » et de maîtrise d'une forme discursive, ce sont aussi celles que soulève la circulation des concepts d'une langue à l'autre, d'une discipline à l'autre ; en d'autres termes, il s'agit de la question du sens et de la capacité à le partager au-delà des barrières linguistiques et disciplinaires.

De façon plus large, depuis les débuts de l'expérience des projets de recherche collaborative dans le cadre des programmes de la communauté européenne, les problèmes soulevés par la pluridisciplinarité, par l'éclectisme des problématiques et de leurs fondements épistémologiques étaient manifestes et assez généralement reconnus. La création du réseau d'excellence Kaléidoscope (2004-2007) dans le cadre du 6^e Programme Cadre de recherche et développement (PCRD) relevait le défi de saisir frontalement ces problèmes. En affichant l'objectif partagé de rechercher la convergence sur les « *Concepts and methods for exploring the future of learning with digital technologies* », ce réseau affirmait un objectif ambitieux : « *The role of Kaleidoscope is to overcome this complexity by addressing it directly in search of a common ground, a shared theoretical and methodological framework beyond disciplinary specificities, but clearly and firmly rooted in the disciplines.* » (Collectif, 2003, p. 6).

Les nombreuses équipes du réseau appartenait à des traditions et des organisations de la recherche très diverses. Nombre d'entre elles n'avaient pas encore eu d'occasions de collaborer. Aussi la priorité a-t-elle été donnée, lors du lancement du réseau, à la création de relations scientifiques par le moyen de projets communs limités, mais concrets. Les efforts et l'énergie nécessaires à cette construction ne permirent pas d'engager d'emblée le travail au cœur de l'objectif fondateur¹.

Coordinateur du réseau, constatant la difficulté à susciter les échanges aux niveaux théorique et méthodologique, j'ai pris l'initiative de mettre l'accent sur une approche pragmatique partant des textes des publications pour identifier les mots et expressions de la recherche, faire leur inventaire et poser la question de *leurs définitions*, c'est-à-dire questionner la nature des concepts auxquels elles renvoient. Il s'agissait de mettre en évidence la diversité des significations selon les disciplines et les cultures scientifiques, et celle de la richesse lexicale spécifique de la recherche sur les « *Technology Enhanced Learning* » (TEL). Ainsi pourrait-on toucher du doigt les points de convergence et de divergence entre les problématiques, les concepts, les méthodes et les pratiques en incluant la complexité liée à la diversité des disciplines et des langues. La constitution de ce lexique multidisciplinaire et international n'était donc pas une fin en soi mais une proposition tactique, dont la visée était d'initier la construction d'un socle scientifique de référence à la fois explicite, divers et collectif.

La première pierre de ce projet fut la constitution d'un entrepôt de publications, disponibles en libre accès, issues des travaux des équipes membres du réseau Kaléidoscope. La deuxième phase fut de réaliser l'inventaire des termes de ce corpus et de leurs relations telles que leur coprésence dans les textes les suggère. La phase finale a consisté en la rédaction d'un dictionnaire pour un échantillon de termes les plus significatifs, en laissant à un projet futur la tâche de le compléter. Cette dernière phase, visant à apporter « une preuve de concept », a été pour une large part réalisée dans le cadre du réseau d'excellence STELLAR du 7^e PCRD. Le produit encore inachevé est le « dictionnaire EIAH » accessible sur le site de l'Association des Technologies de l'Information pour l'Éducation et la Formation (ATIEF, <http://atief.fr/>).

¹ Il y eut cependant des exceptions notables, comme le projet TELMA dont l'objectif était de relier et d'intégrer les cadres théoriques utilisés par différentes équipes dans le domaine des EIAH pour l'apprentissage des mathématiques (Artigue *et al.* 2009).

La suite de cet article présente les réalisations de ces différentes phases. Au préalable, les sections 2 et 3 abordent respectivement le cas particulier et emblématique du nom du domaine, puis le problème de la définition, affirmé comme l'un des requis du développement scientifique. Les sections 4 et 5 décrivent la construction du corpus des termes, le prototype du thésaurus et enfin la construction du dictionnaire.

2. Jalons d'une histoire, le nom du domaine

Le nom d'un domaine de recherche tient une place bien particulière dans son vocabulaire, il est une métadonnée indispensable à son identification dans le champ scientifique et à celle de ses produits, notamment les publications. Son histoire reflète ses évolutions internes comme celles de ses relations avec son environnement institutionnel et social. Ainsi, l'expression « *Environnements informatiques pour l'apprentissage humain* » (EIAH), aujourd'hui largement utilisée en France, a-t-elle succédé à d'autres désignations pour refléter l'évolution de son contexte, notamment technologique, ou pour des besoins de structuration interne. C'est un peu de cette histoire qui est décrite dans les lignes qui suivent, en précisant les principaux jalons qui illustrent la complexité et la diversité des influences qui orientent l'évolution de la recherche.

Le nom du domaine dans lequel nous nous inscrivons aujourd'hui a son origine dans la restructuration de la recherche en informatique et mathématiques appliquées à Grenoble. Dans ce contexte, l'enjeu pour l'équipe, qui avait alors pour nom « Didactique et technologies cognitives en mathématiques », était d'affirmer clairement l'articulation de la recherche en informatique, alors essentiellement représentée par le projet Cabri-géomètre (Laborde, 2003), et de la recherche sur l'enseignement et l'apprentissage, alors essentiellement représentée par la didactique des mathématiques. Le choix s'est porté sur « *Environnements informatiques d'apprentissage humain* »², permettant une meilleure lisibilité de l'orientation scientifique pour la communauté informatique sans pour autant se détacher des sciences cognitives en plein développement à l'époque. Le qualificatif « humain » permettait en outre de se distinguer de l'appren-

2 Le problème du projet scientifique de l'équipe dans la communauté informatique a été posé à la suite de son audition par la Commission scientifique de l'institut IMAG en octobre 1993. Le nouveau nom et le projet associé ont été validés en 1995 dans le cadre de la restructuration quadriennale de l'institut. Pour la période quadriennale 1995-1999, l'équipe EIAH a rejoint le Laboratoire Leibniz fondé par Philippe Jorrand.

tissage machine. Quant à « environnement » il permettait une ouverture assez large pour inclure des projets logiciels, d'ingénierie didactique ou des problématiques d'enseignement à distance. L'introduction de la préposition « pour » est apparue un peu plus tard, sur une suggestion de Monique Baron, corrigeant la formulation initiale maladroite. Outre un équilibre renforcé entre les disciplines, cette préposition fait clairement apparaître le problème scientifique : créer les connaissances et les méthodes pour concevoir, réaliser et utiliser des environnements informatiques pour susciter, accompagner et évaluer des apprentissages humains. L'expression « *Environnements informatiques pour l'apprentissage humain* » est reprise en 1997 dans une contribution collective au colloque PRC IA de Grenoble (Balacheff, Baron, Desmoulins, Grandbastien et Vivet, 1997).

Dans les années 90, en France, la recherche était organisée en deux grands pôles. D'une part, la communauté « *Environnements interactifs d'apprentissage avec ordinateur* » (EIAO)³ et, d'autre part, la communauté « *Hypermédiat et apprentissages* ». Des chercheurs en informatique, éducation, psychologie et didactique participaient aux deux communautés avec une coloration informatique plus forte du côté de l'EIAO, et une coloration recherche en éducation plus forte du côté de « *Hypermédiat et apprentissages* ». La séparation de ces deux communautés avait pour conséquence une dispersion des forces qui affaiblissait la lisibilité du domaine pour les institutions. Le rapprochement s'est concrétisé, à l'initiative de l'ATIEF, par la fusion de la conférence « *Hypermédiat et Apprentissages* » et des « *Journées Environnements Interactifs d'Apprentissage avec Ordinateur* » en une conférence commune, « *Environnements Informatiques pour l'Apprentissage Humain* », dont la première édition a été organisée à Strasbourg en 2003 (<http://archiveseiah.univ-lemans.fr/EIAH2003>). L'ambition était clairement fondatrice :

« Cette dynamique de rassemblement scientifique porte une ambition qui va plus loin que la fusion des participants et des problématiques. Avec cette nouvelle conférence, construite sur les expériences et les acquis des conférences d'origine, il s'agit tout d'abord de bâtir les contours d'un domaine nécessitant des recherches spécifiques que ne peuvent conduire chacune des disciplines qui contribuent à ses travaux (didactique des

3 Cette déclinaison du sigle EIAO date de 1990 (Baron *et al.*, 1991). Elle succède à « *Enseignement intelligemment assisté par ordinateur* » qui désignait un courant de recherche en IA, né dans les années 80 dans le sillage des travaux sur la représentation des connaissances (Baron, 1994).

disciplines, informatique, psychologie cognitive, sciences de l'éducation, sciences de la communication, etc.)» (Desmoulins, Marquet et Bouhineau, 2003).

L'expression « *Environnements informatiques pour l'apprentissage humain* » n'a pas de correspondance immédiate parmi les désignations du domaine en anglais. Bien que la traduction « *Informatic environments for human learning* » soit correcte, son adoption par la communauté internationale est peu probable. Un équivalent possible serait « *Computer-based learning environment* » dont la signification littérale a quelque proximité avec celle d'EIAH : « *[it] refers to a broad array of uses of technology that are aligned with theories that support learning* » (Lajoie et Naismith, 2011, p. 716) ; cependant, l'accent mis sur l'usage laisse de côté la part de recherche informatique ou technologique propre au domaine. En fait, comme pour les autres termes de la recherche, il s'agit moins de traduire que de trouver dans le vocabulaire anglo-saxon les désignations qui correspondraient le mieux à EIAH ; en tout cas, il est nécessaire de trouver un appariement meilleur que celui proposé par Wikipedia qui, pour son équivalent en anglais, redirige la page « Environnements informatiques pour l'apprentissage humain » (https://fr.wikipedia.org/wiki/Environnements_informatiques_pour_l%27apprentissage_humain) vers la page « *Intelligent Tutoring Systems* »⁴.

Les dénominations dominantes dans le monde anglophone sont : « *Educational Technology* » ou « *Instructional Technology* » aux États-Unis et en Australie, « *Learning Technology* » en Grande-Bretagne et « *Technology Enhanced Learning* » en Europe. Ces expressions ont une signification assez large, il en existe d'autres pour des thématiques plus spécifiques (par ex. *Artificial Intelligence in Education* - AIED, *Intelligent Tutoring Systems* - ITS, *Learning Analytics and Knowledge* - LAK, *Computer supported collaborative learning* - CSCL, *Mobile Learning* - ML).

La plus ancienne association aux États-Unis, « *Association for Educational Communication and Technology* » (AECT), a une tradition de veille sur la terminologie du domaine et notamment reconsidère régulièrement sa définition et son nom. Ainsi a-t-elle décidé en 2008 d'abandonner la dénomination « *Instructional technology* » pour revenir à

⁴ Redirection vérifiée le 6 juillet 2018 à 17:18.

« *Educational Technology* » (Hlynka, 2013, p. 103). Cette décision ne fait pas l'unanimité (Lowenthal et Wilson, 2010). Il est en particulier remarqué que si les deux expressions sont utilisées de façon interchangeable, « *Educational Technology* » renverrait plutôt à l'enseignement (de la maternelle à la fin du lycée) alors que « *Instructional Technology* » serait plus générique en couvrant éducation et formation (Lowenthal et Wilson, 2010, p. 41). Mais surtout, cette dernière apparaît plus fréquemment dans les profils de postes et dans le libellé des cursus universitaires (Lowenthal et Wilson, 2010). La critique du changement de nom ne se fait donc pas essentiellement sur le terrain de la recherche, mais sur celui de la lisibilité pour le monde extérieur. Elle ouvre la discussion sur d'autres possibilités, notamment :

« We are not arguing necessarily for the adoption of one label over another. While there are a growing number of people who recommend adopting a label that includes the word "learning" – such as "Learning Design" – there are others like Reiser and colleagues (e.g., Reiser, 2007) who have for years made a compelling case for the label instructional design and technology. » (Lowenthal et Wilson, 2010, p. 54).

Le terme « *learning* » a été mis en avant dans le discours de la recherche européenne très tôt dans les années 2000. Ainsi, l'expression « *Technology Enhanced Learning* » (TEL), utilisée communément en Europe, est apparue lors du lancement du 6^e PCRD de la Communauté européenne. Le document support de la journée d'information sur ce programme, le 5 décembre 2002 à Luxembourg, l'introduit discrètement :

« Under the 6th RTD Framework program (FP6), the Technology Enhanced Learning (e-learning) has been set as strategic objective in the thematic priority "Information Society Technologies" of the Specific Program "Integrating and Strengthening the European Research Area". The text of the FP6 specifies that "Work on eLearning will focus on personalised access to, and delivery of, eLearning as well as on advanced learning environments at school, university and in the workplaces that take advantage of the development of ambient intelligence". » (European Commission 2002, p. 32).

TEL apparaît donc comme un synonyme de *e-learning* (ou *eLearning*). Pourtant, l'objectif de la nouvelle dénomination est de donner un signe politique fort avant d'être scientifique. Il s'agit d'infléchir la recherche dans une direction mettant en avant l'apprentissage pour des décideurs politiques peu convaincus par les avancées du *e-learning* (trop focalisées

sur les ressources). La recherche « centrée apprenant » était alors une injonction en écho au mot d'ordre de la fin des années 90 : « mettre l'élève au centre du système éducatif ». À la fin du 6^e PCRD, le nouveau nom du domaine est largement adopté. Sa signification n'est pas très précise, mais se cherche dans une direction cohérente avec l'orientation à l'origine du consensus en France sur l'expression EIAH. Ainsi Pat Manson⁵ affirme-t-elle en 2007 :

« Technology-enhanced learning may not flow readily off the tongue or be easily translated as a brand name, but it very consciously reflects what it is: Using ICT to secure advancement in learning. [...] Technology Enhanced Learning combines but places equal emphasis on all three elements - on technologies, on learning and on enhancement or improvements in learning. » (Manson, 2007).

« *Technology Enhanced Learning* » n'a donc pas son origine dans une réflexion de la communauté scientifique sur sa propre évolution, mais dans la volonté politique de la Commission européenne d'infléchir l'orientation de la recherche et de l'innovation. L'accent mis sur l'apprentissage (*learning*) est toutefois cohérent avec l'évolution du domaine et avec le choix fait par la communauté française. TEL est ainsi un bon candidat pour un appariement dans le vocabulaire anglophone avec EIAH.

Le choix du nom du domaine, c'est-à-dire la clarification et l'affirmation de son identité, est un acte fondateur difficile, complexe et souvent source de tension. Il est nécessaire d'une part pour situer le domaine dans l'ensemble du champ scientifique et relativement aux organisations qui le structurent. Il est aussi nécessaire à sa structuration interne. Plus largement, ce travail de définition et de dénomination doit être entrepris pour l'ensemble des concepts que la recherche forge, souvent de façon originale, pour avancer dans son travail de connaissance.

⁵ Directrice de l'unité « *Cultural Heritage and Technology Enhanced Learning* » de la direction générale « *Information Society and Media* » (Commission Européenne).

3. Définir, un acte nécessaire au développement scientifique

3.1. Complexité et enjeux de la langue

La recherche sur les EIAH implique des prises de positions épistémologiques, théoriques ou philosophiques marquant un rapport aux savoirs, une conception de l'apprentissage voire, plus généralement, une vision de l'éducation. Qu'ils soient explicites ou non, conscients ou non, ces positionnements se manifestent lors de la présentation des problématiques ou de l'argumentation pour affirmer l'intérêt des approches et de leurs résultats. Ainsi la maîtrise de la langue naturelle et, dans une certaine mesure, celle de la rhétorique jouent-elles un rôle important dans ce domaine qui, au moins dans son état actuel, ne dispose pas de cadres théoriques ou de modèles assez formalisés ni de méthodes suffisamment normalisées. Par ailleurs, poussée par la technologie et tirée par le marché ou la demande des politiques publiques, la recherche doit intégrer des courants de pensée et développer des stratégies d'adaptation rapide, qui se traduisent notamment par des évolutions et une instabilité du vocabulaire qui rendent plus fragile la construction scientifique. L'irruption des MOOC, produit technologique et innovation institutionnelle (voire pédagogique), en est un exemple particulièrement évident. Elle a, par exemple, propulsé sur le devant de la scène médiatique et scientifique les mots clés « connectivisme », « *learning analytics* », « classe inversée », et suscité un effet de priorité dans les choix et les soutiens institutionnels.

Les mots clés sont des marqueurs significatifs de l'histoire de la recherche sur les EIAH. Des mots tels que « tuteur », « micromonde », « *e-learning* », « jeux sérieux » ou « MOOC » se sont imposés à diverses époques pour exprimer l'orientation de programmes scientifiques nés sous des impulsions diverses qui vont des avancées de l'informatique – l'intelligence artificielle – à des espoirs économiques – l'émergence d'un modèle de formation massif et ouvert. Des thèmes de recherche sont renouvelés au fil de ces évolutions, comme celui de la prise en compte de l'apprenant auquel renvoient successivement des expressions telles que « modélisation de l'apprenant », « fouille de données éducatives », « *learning analytics* ». Chacune marque une évolution, sans cependant être accompagnée de la disparition de ce qui existait jusque-là : « modélisation de l'apprenant » s'appuyait sur des modèles que l'on pourrait qualifier de psychologiques, « fouille de données éducatives » reflétait la

reconnaissance du progrès des techniques d'analyse exploratoire des données, enfin « *learning analytics* » importe dans le champ de l'éducation des techniques d'analyse de grands ensembles de données (corpus d'usage, traces massives) qui font leurs preuves ailleurs. Dans ces deux derniers cas, les relations entre le nouveau et l'ancien terme ne sont pas clairement précisées. Alors que la « fouille de données éducatives » se trouve dans le fil du développement de la recherche sur les tuteurs intelligents (Bourdeau et Grandbastien, 2011, p. 4), les « *learning analytics* » marquent l'impact sur la recherche en EIAH d'avancées dans des domaines assez éloignés, mais centrés sur l'utilisateur humain, tel le commerce électronique (Balacheff et Lund, 2012).

Un très grand nombre de termes, ont été forgés dans le contexte de la recherche anglophone. Certains ont été traduits, certains ne l'ont pas été, d'autres enfin ont été importés mais cohabitent avec une traduction. Ainsi « *educational data mining* » est-il traduit par « fouille de données éducatives », thème par exemple de la Journée EIAH & IA 2015 (<http://pfia2015.inria.fr/journees-bilaterales/journee-eiah-ia>). Au contraire, la traduction de « *overlay* », terme qui désigne une approche particulière de la modélisation de l'apprenant dans les « tuteurs intelligents » (Carr et Goldstein, 1977), par « expertise partielle » proposée par Vivet et Nicaud (1988), bien qu'elle soit fidèle à la définition originelle, ne sera pas adoptée. En revanche, « *serious game* » et sa traduction « jeux sérieux » cohabitent dans le vocabulaire francophone comme en atteste le sommaire du numéro spécial « Évaluation dans les jeux sérieux » de la revue STICEF (volume 21, 2014, http://sticf.univ-lemans.fr/num/vol2014/sticf_2014_som.htm#evajs). Cette cohabitation semble relever d'une tendance assez fréquente de préférer l'anglicisme sans qu'il y ait de justification sémantique. Il n'en va pas de même de la cohabitation dans l'usage du terme « *e-learning* » avec sa traduction « formation en ligne » entrée dans la terminologie française en 2005⁶ (Journal officiel, 2005). L'origine de cette cohabitation serait plutôt à rechercher dans l'acception très générale de l'expression « formation en ligne » qui inclut des catégories difficilement acceptées sous le terme *e-learning*, par exemple celle des MOOC. Cette situation soulève la question de ce que serait une définition

6 Journal officiel du 14/05/2005 : « Définition - Formation faisant appel à des moyens de communication électroniques. Note: Le terme "e-formation" est à éviter. » Le français canadien a retenu « apprentissage en ligne » (Pavel 2002).

précise de « *e-learning* » qui distinguerait ce terme de celui de « formation en ligne ». L'analyse de la littérature anglophone (Moore *et al.*, 2011) atteste à la fois de la grande variété de définitions disponibles et de l'impossibilité de donner une définition les incluant toutes qui ne serait pas finalement trop générale (Sangrà, Vlachopoulos et Cabrera, 2012). Le terme *e-learning* est donc polysémique, ce qui est rare dans la pratique scientifique, mais peut être toléré si à une définition particulière peut être associée une problématique précise et spécifique.

Le problème de la traduction n'a véritablement d'importance que si l'origine, la raison d'être et le choix des termes jouent un rôle effectif dans la pratique scientifique. Un exemple en mathématique, quoiqu'anecdotique, est éclairant : le terme « corps » est utilisé en français pour désigner une structure algébrique forte dont la traduction pour les mathématiciens anglophones n'est pas « *body* » mais « *field* », écart étrange que seule explique la « petite » histoire du concept et celle de sa dénomination⁷. Mais cet écart importe peu dans la pratique ; en effet, le couplage explicite, technique et précis entre termes et définitions permet en mathématique un appariement rigoureux entre les termes des différentes langues. Il n'en va à l'évidence pas de même pour la recherche en EIAH. Le locuteur francophone prend particulièrement conscience de cette situation lorsqu'il souhaite s'exprimer dans une autre langue que la sienne, le plus souvent l'anglais. Cela vaut *a priori* pour d'autres langues sources, cependant toutes n'ont pas la même politique. Les universitaires hollandais, par exemple, considèrent l'anglais comme leur langue (naturelle) de travail, ainsi cette langue est-elle utilisée pour la rédaction des thèses.

Si l'anglais est devenu la *lingua franca* du domaine, il ne s'agit pas – ou pas simplement – de la conséquence d'une forme d'impérialisme comme on le suggère parfois. Trois autres raisons doivent être évoquées : (i) l'antériorité et le nombre significatif des travaux sur le continent

7 « C'est [Dedekind] qui introduit le mot *corps* (*Körper* en allemand, d'où la notation K toujours très utilisée) pour un ensemble où, à l'instar des rationnels ou des réels, on a une addition, une multiplication et une division. Ce nom est motivé par la richesse de cette structure. L'intention de Dedekind est d'évoquer un organisme vivant. Les anglo-saxons, les français et les russes traduisent *Körper* par le mot passe-partout de *champ*. Mais les francophones finissent par adopter le mot *corps* (est-ce l'influence de Bourbaki, très marqué à ses débuts par l'école allemande ?), alors que jusqu'aujourd'hui, "*field*" subsiste en anglais, "*polié*" ["Поле "] en russe » (Lafontaine, 2011).

américain - voir à ce sujet (Bruillard, 1997), (ii) la productivité de la recherche anglo-saxonne dans le sillage du développement académique, industriel et économique de l'informatique et en particulier de l'intelligence artificielle, (iii) le contexte institutionnel et professionnel de l'enseignement ou la formation plus favorable aux relations entre recherche, industrie et usages.

L'enjeu de la langue n'est pas qu'une question de mots, comme on l'entend trop souvent affirmé, mais celui de l'expression et de la circulation du sens. C'est un enjeu scientifique car seules les problématiques, les modèles et leurs fondements théoriques précisément explicités et les mots précisément définis permettent le progrès de la connaissance par l'examen critique des résultats, le débat sur leur validité et finalement leur indispensable capitalisation. La nature et l'importance scientifique des problèmes de langue sont révélées lors de l'expression dans une langue différente de la langue de travail du locuteur, notamment à des fins de publication. C'est donc aussi un enjeu de politique scientifique parce que la maîtrise de la langue est un élément déterminant de l'intégration dans la communauté scientifique. On peut pourtant relever le peu d'appétence pour traiter ces problèmes – et ce n'est pas une spécificité française ou francophone – comme en témoigne cette conclusion d'une étude sur l'importance des différences de signification entre les termes « *e-Learning* », « *online learning* » et « *distance learning environments* » :

« The findings show great differences in the meaning of foundational terms that are used in the field, but also provide implications internationally for the referencing, sharing, and the collaboration of results detailed in varying research studies. » [...] « The lack of consistency in terminology inevitably affects not only the researchers who would like to build upon the findings, but also impacts designers who are creating similar types of environments. Terminology also poses a problem when the specific context of the learning environment is not described in sufficient detail or when its identification is not very prominent in both the discussion of the methods and the other sections of the paper. » (Moore, Dickson-Deane et Galyen, 2011, p. 134).

3.2. Note sur la réticence à définir

L'idée de s'astreindre à définir les termes du vocabulaire de la recherche sur les EIAH et de les utiliser, sauf à justifier une évolution ou un rejet, est considérée avec réticence par de nombreux chercheurs. Cette

défiance reflète la perception d'un risque attaché au travail définitoire, ou plus largement à la formalisation des connaissances et des théories. Cela ne signifie pas que ce travail soit ignoré, mais il n'est conçu que dans le cadre local de projets spécifiques et non comme une entreprise collective qui lierait les chercheurs entre eux en créant une obligation d'usage et de référence. Il est tout à fait plausible qu'il s'agisse là d'une réaction protectrice pour échapper au dogmatisme ou à l'orthodoxie dont Edgard Morin (1991, p. 129-sqq) rappelle qu'elles sont des tendances humaines naturelles. Il est vrai que l'usage de certaines références prend parfois une allure d'allégeance à une doxa; ce fut largement le cas pour le *constructivisme* qui a dominé la recherche au siècle dernier, le plus souvent pour invoquer une perspective ou un cadre général de pensée (l'apprenant construit sa connaissance) sans entrer dans la complexité de ses concepts spécifiques et de ses modèles.

Moins qu'une défiance, ce peut être un manque d'intérêt qui tient au grand écart que doivent gérer les projets entre recherche scientifique et innovation technologique. En effet, la recherche sur les EIAH a les caractéristiques d'une recherche fondamentale qui a pour but de contribuer à accroître la connaissance sur l'apprentissage et l'enseignement (ou la formation) et celles d'une recherche « visionnaire » qui explore les voies qui permettront de produire des technologies innovantes pour apprendre et enseigner plus efficacement. Il y a dans cette activité une part *d'invention* dont les sources sont des idées, des intuitions, des convictions aux origines multiples soutenues par des arguments dont la force tient moins à la rigueur des méthodes et à la solidité des théories qu'à la réussite revendiquée des réalisations. C'est d'ailleurs cette invitation à l'innovation qui domine le discours accompagnant les appels à projets européens qui, pour une grande part, tirent la recherche par les opportunités de financements importants qu'ils offrent. Il en va de même pour les appels à projet de l'Agence nationale de la recherche (ANR). La politique de financement de ces institutions pousse la recherche en EIAH au-delà de la distinction classique entre recherche fondamentale et recherche appliquée, vers la recherche-développement (R&D). Cette compétition, alors que les fondements et les outils sont peu assurés, a pour effet de favoriser les courants et les modes accompagnés de la création rapide de journaux et workshops avec leurs vocabulaires propres comme marques distinctives. La formalisation et l'approfondissement des cadres théoriques et des modèles apparaissent, dans ce contexte, moins importants que des indicateurs de l'efficacité potentielle des technologies proposées

et ceux de la rapidité avec laquelle elles « vont » sur le marché. Au fond, comme le remarque Moore à propos de la définition de *e-learning* :

« What is abundantly obvious is that there is some uncertainty as to what exactly are the characteristics of the term, but what is clear is that all forms of e-Learning, whether they be as applications, programs, objects, websites, etc., can eventually provide a learning opportunity for individuals. » (Moore et al., 2011, p. 130).

En d'autres termes, ce qui importe n'est pas le discours qui accompagne les technologies d'apprentissage, mais l'évidence de leur utilité. On peut concevoir que cela suffise lorsque le discours a pour principale fonction de permettre de situer un projet, de le relier à d'autres, ou de convaincre de son caractère innovant ou de son efficacité. Ce ne l'est plus lorsque le discours a une valeur d'usage et non d'échange, c'est-à-dire lorsqu'il permet l'analyse critique, le jugement de validité, la reprise des concepts ou des méthodes, qu'il s'agisse de la conception, de la réalisation ou des usages. Les mots sont alors des outils. Leurs définitions sont nécessaires pour, en quelque sorte, en objectiver le sens et le discuter en référence à leurs utilisations dans la construction de modèles, de théories ou dans la prétention de résultats. Cette explicitation est indispensable pour mettre en relation les travaux, examiner la réalité de progrès et constituer un socle sur lequel la recherche peut avancer. Dans la construction de la connaissance une première étape est de nommer, sans quoi toute communication scientifique est impossible, puis il faut au fil des échanges s'assurer de la portée des accords ou désaccords et pour cela expliciter le sens, c'est-à-dire définir. Pour autant, les définitions ne sont pas des objets figés ou intouchables. Elles seront travaillées pour être précisées, reformulées. De nouveaux termes seront éventuellement forgés. Le consensus sur les définitions doit s'entendre comme un état provisoire ; il ne referme pas mais prépare l'approfondissement qu'initiera la vigilance critique qui conduira à des remises en question, des évolutions, refontes ou abandons.

4. Le corpus et le thésaurus des termes et expressions de la recherche sur les EIAH

Cette section et la suivante présentent la construction, respectivement, du thésaurus et du dictionnaire, leurs structures et leurs organisations éditoriales.

4.1. La construction d'un corpus, les archives ouvertes TeLearn

Les archives ouvertes, dans les années 2004-2007, étaient considérées avec réserve par les chercheurs et les équipes partenaires du réseau européen Kaléidoscope, notamment par ceux appartenant aux communautés de Sciences humaines et sociales. Le statut juridique du libre accès leur paraissait incertain au regard des droits d'auteur ou d'éditeur. Par ailleurs, le risque d'usage abusif, celui de voisiner avec des publications médiocres et l'attachement aux formes classiques de publication chez des éditeurs historiquement reconnus freinaient l'adoption de ce mode nouveau de diffusion. Pour sa part, la France avait en Europe une relative avance avec la création, sur la plate-forme Hyper-Article en Ligne (HAL) du Centre pour la communication scientifique directe (CCSD), du portail EIAH à l'initiative du RTP39, réseau thématique pluridisciplinaire du CNRS et du portail EduTICE dans le cadre du programme expérimental TemaTice du Ministère de la recherche. Mais les outils proposés à l'époque par le CCSD ne permettaient pas de mettre en place une exploration fine du corpus, ni d'identifier les sources avec la souplesse que nous souhaitons dans le cadre de la construction du réseau (notamment l'utilisation de mots clés standards éventuellement complétés par des mots clés libres).

La solution fut de créer une instance privée de HAL, nommée TeLearn (Min, Balacheff et Zeiliger, 2008), en assurant la compatibilité nécessaire pour pousser dans l'entrepôt du CCSD les documents réunis dans l'archive locale. Cette instance fut augmentée des fonctionnalités nécessaires au projet (que nous ne détaillerons pas ici). La première présentation publique de l'archive a eu lieu lors de la conférence IST2006 à Helsinki. Très favorablement reçue, la nouvelle fut largement relayée par la Commission européenne. Le portail TeLearn (<https://telearn.archives-ouvertes.fr/>) a été intégré à la plate-forme HAL en 2010 dans le cadre du réseau Stellar.

4.2. La constitution du vocabulaire

La liste des termes du vocabulaire est le résultat d'une analyse lexicale limitée aux publications de langue anglaise alors disponibles dans l'archive TeLearn, soit 1068 fichiers, avec le logiciel AntConc (Anthony, 2010); le choix de l'anglais tient aux raisons évoquées plus haut (sous-section 2.1).

AntConc est une boîte à outils sous licence logiciel libre (Anthony, 2004) qui permet de constituer assez simplement un lexique (termes simples et composés) à partir d'un corpus textuel en relevant notamment le nombre des occurrences. Sa capacité de traitement le limite à de petits volumes, mais cela ne constituait pas un obstacle à cette étape du projet.

Trois ensembles de termes sont distingués : d'une part, les termes disciplinaires généraux (par exemple : « schème » en psychologie, « système multi-agents » en informatique, « professeur » en pédagogie), d'autre part, les termes disciplinaires ayant un rôle significatif dans la recherche sur les EIAH (par exemple : « conflit sociocognitif » en psychologie, « environnement auteur » en informatique, « évaluation formative » en pédagogie) et enfin les termes spécifiques de la recherche sur les EIAH (par exemple : « script de collaboration », « système de gestion des apprentissages », « apprentissage nomade »). La priorité a été donnée à ce dernier ensemble, avec le projet, dans une seconde phase, d'étendre la liste aux termes importés des diverses disciplines qui ont une contribution originale ou spécifique. Les termes du premier ensemble ne sont considérés que dans la mesure où ils sont d'un usage fréquent et potentiellement problématique pour les chercheurs n'ayant pas de formation dans la discipline dont ils sont issus.

Cette analyse a produit un ensemble de 611 termes simples, retenus au seuil de 1000 occurrences, et un ensemble de 729 expressions au seuil de 100 occurrences. Ces seuils respectifs de 1000 et 100 occurrences ont été déterminés empiriquement, en vérifiant qu'ils n'écartaient pas de termes importants pour le domaine⁸ ; les 10 termes et les 10 expressions de plus haute fréquence sont présentés dans le tableau 1.

Les listes obtenues ont été analysées pour en retirer les termes trop génériques (par exemple : « *email* », « *website* », « *research* ») ou trop spécifiques tels que les noms propres (par exemple : « *Algebrista* », « *European Commission* »). Le résultat de ce désherbage a été la constitution d'une liste de 471 termes parmi lesquels on retrouve 99 mots clés proposés par les auteurs au moment du dépôt de leurs publications.

⁸ Ces réglages empiriques, dépendant de l'expertise des personnes qui conduisent l'analyse, est une limite classique de la construction d'un vocabulaire (Aussenac-Gilles *et al.*, 2000).

Tableau 1 · Termes et expressions de plus haute fréquence dans l'archive TeLearn

Termes de haute fréquence		Expressions de haute fréquence	
104974	learning	6903	learning environments
58379	students	5246	e-learning
37548	knowledge	5174	collaborative learning
31379	system	4340	problem solving
28488	learners	3647	learning process
27141	technology	2905	higher education
25874	computer	2580	learning activities
25301	teachers	2514	teaching and learning
23164	environment	2015	networked learning
23048	research	1789	knowledge construction

Ce vocabulaire a été le matériau à partir duquel ont été construits le thésaurus et le dictionnaire présentés ci-après.

4.3. Le thésaurus

L'intérêt de la construction du thésaurus est de constituer le réseau des relations entre les termes du vocabulaire du domaine et de préciser les concepts associés. À cela nous avons ajouté l'objectif d'établir les relations entre les lexiques dans différentes langues (anglais, arabe, bulgare, chinois classique, chinois simplifié, coréen, danois, espagnol, estonien, français, grec, hollandais, hongrois, italien, japonais, portugais, russe, slovaque, turc, vietnamien).

Cette construction a été réalisée avec le logiciel libre TemaTres (<http://www.vocabularyserver.com/>), retenu parce qu'il permet la création de thésaurus multilingues et multi-hiérarchiques. Ce logiciel a de plus la caractéristique d'être multi-utilisateurs, permettant une construction collaborative. Enfin, il respecte les standards SKOS (<http://www.w3.org/2004/02/skos/>), Dublin-Core et RDF - dans l'éventualité d'un passage à une ontologie.

L'ambition de disposer d'un thésaurus multilingue (figure 1) a posé le problème de la normalisation des termes, qui n'a pas de réponse simple car les normes sont différentes selon les langues. Ainsi, pour la langue anglaise, selon la norme ISO, une distinction doit être faite entre ce que

L'on peut dénombrer (*how many ?*), qui requiert le pluriel, et ce qui ne peut l'être (*how much ?*), qui requiert le singulier. En français, au contraire, le singulier est privilégié, sauf pour les termes utilisés naturellement au pluriel. Nous avons respecté au mieux cette norme pour ces deux langues.

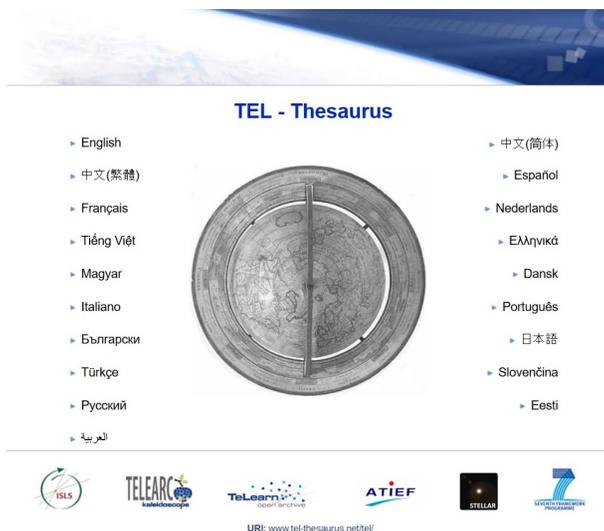


Figure 1 • Page d'accueil du thésaurus

La hiérarchisation du thésaurus rend compte de plusieurs types de relations.

- Les *relations hiérarchiques* formalisent un rapport de subordination entre les notions et facilitent une navigation verticale : passage du général (*broader term*, noté BT) au spécifique (*narrower term*, noté NT).

- Les *relations d'équivalence* relient les synonymes et quasi-synonymes : l'un des termes est préféré (le descripteur), les autres équivalents ne sont alors pas retenus pour l'indexation (les non-descripteurs); nous avons eu recours à ce type de relation pour le thésaurus danois (*i.e.* « *udvidet læringsmiljø* » comme descripteur et « *berigende læringsmiljø* » comme non-descripteur).

- Les *relations associatives* marquent les analogies entre des termes ayant une proximité sémantique (*related term*, noté RT).

Les *relations inter-linguistiques* qui associent les thésaurus des diverses langues sont traitées comme des relations d'équivalence. Nous avons retenu le thésaurus anglais pour pivot de la construction multilingue, pour les raisons évoquées ci-dessus. Il contient donc tous les termes

retenus, les autres thésaurus contiennent la traduction ou les termes appariés (voir un exemple en figure 2). Cette approche n'exclut pas la prise en compte, dans les divers thésaurus, de termes qui n'auraient pas de traduction en anglais, mais auraient une existence liée à l'originalité de travaux conduits dans d'autres contextes linguistiques. Ainsi peut-on trouver en français ou en allemand l'expression « ingénierie didactique », mais pas en anglais. Il est à noter que la mise en relation inter-linguistique n'étant pas symétrique dans TemaTres, les liens réciproques du terme cible vers le terme source sont créés manuellement.



TEL thesaurus - English

Home My account Search

learning management systems (LMS)

Home / learning management systems (LMS)

Term Notes (1) Metadata

learning management systems (LMS)

Related terms

[RT course management systems \(CMS\)](#)
[RT integrated learning systems](#)
[RT personal learning environments \(PLE\)](#)
[RT virtual learning environments](#)

Tiếng Việt hệ quản lí học tập
 Nederlands leermanagementsysteem
 Dansk læringsplatform
 Português sistema de gestão da aprendizagem
 Español sistema de gestión de aprendizaje
 Italiano sistemi per la gestione di corsi
 Français système de gestion des apprentissages
 Magyar tanulás menedzsel rendszer
 Ελληνικά συστήματα διαχείρισης μάθησης
 Български системи за управление на обучението
 中文(简体) 学习管理系统 / 网络教学平台 / 网络教学平台 (cn) 学习管理系统 (sg)
 中文(繁體) 學習管理系統 (tw / hk)

Figure 2 • Exemple d'entrée du thésaurus

Pour des raisons pratiques, la priorité a été donnée aux langues des équipes des réseaux Kaléidoscope puis du réseau Stellar, suivies des langues pour lesquelles existent une communauté d'usage scientifique ou une pratique significative. Pour la plupart des langues, il est rapidement apparu que de nombreux termes n'ont en fait pas de traduction, soit en raison de l'absence de recherches correspondantes dans la communauté considérée, soit en raison de l'utilisation du vocabulaire anglais pour ces recherches.

Le travail de traduction consiste à appairer les termes utilisés par une même communauté internationale ayant une langue de communication et des langues de travail différentes. Cet appariement peut être immédiat parce la langue de travail a adopté un terme de la langue de communication. Ainsi, par exemple, le terme « *e-learning* », lorsqu'il est pris dans une acception technique précise, est-il simplement importé par les locuteurs français ou italien. Dans certains cas des équivalents naturels sont disponibles, par exemple « système auteur » en français ou « *strumento autore* » en italien pour « *authoring tool* », mais l'usage ne les retient pas toujours : la communauté italienne préfère l'expression hybride « *strumento di authoring* ».

En l'absence d'appariement sur la base d'usage attesté, l'exercice de traduction est hasardeux. Par exemple, l'expression « *animated pedagogical agent* » est traduite en français par l'expression « agent pédagogique animé », dont la signification est celle du rapprochement des concepts d'agent pédagogique et d'agent conversationnel animé. En revanche cette expression surprend le traducteur pour le chinois simplifié qui hésite entre deux possibilités : « 动态教学代理 » (agent pédagogique dynamique) et « 动画教学代理 » (agent d'animation pédagogique). Un test minimal sur Google Scholar ne renvoie pas de publications pour la première proposition et seulement une soixantaine pour la seconde ; les deux possibilités seront finalement conservées, laissant la décision à l'utilisateur. En revanche, pour la traduction en chinois classique « 動畫型教學代理人 » (en français : Agents d'animation pédagogiques), Google Scholar ne renvoie pas de publications ; c'est cependant cette expression que le traducteur recommande de retenir. L'expression « *Augmented learning environment* » présente la même difficulté, mais cette fois la traduction proposée « 增量学习环境 » (environnement d'apprentissage incrémental) induit un contresens : en fait, les articles utilisant cette expression portent sur l'apprentissage machine. Dans la même situation, le choix se porte en vietnamien sur « Môi trường học tập cải tiến » (Amélioration de l'environnement d'apprentissage) bien que Google Scholar ne renvoie aucune ressource, laissant ainsi penser qu'il n'y a pas de recherche en vietnamien liée au concept d'environnement d'apprentissage augmenté (à prendre au sens de réalité augmentée) ; l'interprétation littéraire induit un faux sens. C'est en fait une situation fréquente. Ainsi « *buddy system* » n'a pas de traduction en estonien, ou encore « *Agent-based learning environment* » est traduit en italien par « *Ambiente di apprendimento basato sull'uso di agenti* » mais n'est retrouvé dans aucune publication par Google Scholar, ce qui

suggère l'absence d'utilisation de ce concept dans les publications en italien. L'expression « *seamless learning environment* » n'a pas d'équivalent dans le vocabulaire de la recherche française, constat que font aussi les traducteurs hollandais qui envisagent la traduction littérale « *naadloze leeromgeving* », mais observent qu'elle ne renvoie à aucune publication lors de leur recherche de publications (sur internet). La solution peut être de proposer pour traduction une phrase qui explicite le sens de l'expression. Par exemple, en arabe « *على يـبنـي الـذي التـعلم* » (apprentissage qui se fonde sur l'enquête) est proposée pour « *Inquiry learning* ».

Le projet a par ailleurs été l'occasion de mettre en évidence que certaines communautés linguistiques, actives dans le contexte international, n'avaient que peu développé la communication dans leur contexte national; tel est le cas des communautés allemande et hollandaise. Cependant, les traducteurs hollandais, d'abord sceptiques, ont reconnu la richesse du questionnement suscité par cette tâche et ont découvert l'existence d'une communication à l'écart du monde académique. Le même constat a été fait pour le danois.

5. Le dictionnaire

Comme pour toute autre activité collective, les pratiques discursives de la recherche se stabilisent par la force des choses, la collaboration imposant la constitution d'un espace linguistique assez stable et partagé. Au fil du temps, les incertitudes sur les mots sont oubliées alors que paraissent s'imposer des références et des manières de faire communes. Le problème de leur sens ressurgit lorsqu'il s'agit de s'adresser à de nouveaux venus, chercheurs ou étudiants, voulant s'engager dans le domaine. Il se manifeste aussi lorsqu'il s'agit de communiquer avec des interlocuteurs extérieurs à la recherche sur les EIAH, mais utilisateurs potentiels de ses résultats. Il est enfin une circonstance, la formation, dans laquelle il est nécessaire de fixer un lexique et d'explicitier le sens des mots. Dans tous les cas, l'enjeu est celui d'assurer une communication cohérente et stable avec les divers interlocuteurs - qu'ils soient industriels, formateurs ou enseignants, ou bien décideurs du monde de l'enseignement ou de la formation. Le dictionnaire est un instrument au service de cet enjeu.

5.1. Les principes

La création d'un dictionnaire est une tâche complexe dont la réalisation n'a pas de solution canonique. Il faut en premier lieu faire le choix de principes et d'une méthode.

Préciser les sens des termes en explorant la littérature est beaucoup moins simple qu'il n'y paraît. D'une part, peu de textes apportent des réponses explicites et précises et, de plus, ils peuvent révéler une importante variété (p. ex. les nombreuses définitions de « *learning object* »). D'autre part, les dictionnaires sont rares et faiblement documentés. Ils reflètent souvent plus le point de vue d'un auteur qu'un point de vue partagé. La ressource la plus fameuse et fréquemment utilisée, Wikipedia, ne permet pas de combler ce manque en raison de sa conception même. Des points de vue s'y affrontent sans que l'on puisse comprendre les enjeux, ou des définitions sont proposées sans possibilité solide d'évaluer leur validité ou leur singularité. Ainsi, l'expression « *learning design* », qui est attachée à un ensemble de chercheurs assez large, se voit-elle renvoyée pour sa définition à « *instructional design* », créant ainsi une relation de dépendance conceptuelle tout à fait discutable (https://en.wikipedia.org/w/index.php?title=Learning_design&redirect=no⁹). Par ailleurs certains termes ont des significations différentes mais aussi des proximités dont il faut rendre compte, c'est le cas par exemple des termes « *scenario* », « *script* » et « *orchestration* ». Le terme « *script* », notamment, utilisé en psychologie et en informatique, a un sens particulier qui s'est forgé dans le cadre des recherches sur le CSCL (« *collaboration script* », « *external script* », « *internal script* ») et qu'il est important de saisir pour éviter les contresens. En effet, le terme « *script* » n'a pas, dans la recherche en EIAH, le sens donné en programmation. Il est importé de l'IA où il signifie une séquence d'actions qui caractérise une situation bien définie, mais sa signification est proche de celle forgée par la psychologie cognitive. La recherche sur les CSCL a introduit dans sa définition la possibilité de degrés de liberté de l'apprenant (http://www.tel-thesaurus.net/wiki/index.php/Collaboration_script).

Par ailleurs, dans un domaine dont les cadres théoriques propres sont peu formalisés et peu stabilisés, et dans lequel les emprunts à diverses

9 Vérifié à 18:19 le 28 décembre 2017.

disciplines sont importants, le travail de définition apporte une contribution à la conceptualisation et à la différenciation des notions clés. C'est ainsi le cas de l'expression « *educational affordance* » forgée pour préciser ce que le concept d'affordance, très présent dans les recherches mettant au premier plan les interactions personnes-systèmes, a de spécifique et de productif dans un contexte d'apprentissage. Les affordances éducatives peuvent être définies comme les relations entre les propriétés d'une intervention éducative et les caractéristiques de l'apprenant qui permettent la réalisation de certains d'apprentissage (http://www.tel-thesaurus.net/wiki/index.php/Educational_affordance/fr). C'est aussi le cas de l'expression « *epistemic feedback* » dont la définition permet de distinguer explicitement les feedbacks qui jouent un rôle dans la construction du sens d'une connaissance ou d'une compétence. Ce concept a été introduit dans le cadre de la conception et de l'étude de réalités virtuelles et de simulations pour la formation professionnelle, pour caractériser les feedbacks qui permettent à l'apprenant d'analyser son action en relation avec la connaissance en jeu (http://www.tel-thesaurus.net/wiki/index.php/Epistemic_feedback).

Dans ce contexte, la construction du dictionnaire EIAH a pris pour principe de partir de définitions publiées dans des revues ou des actes de conférence, avec la double contrainte de retrouver les racines historiques et de rendre compte de leur diversité éventuelle et de leurs évolutions. Chaque définition proposée est consolidée par une documentation permettant de préciser, chaque fois que cela est possible, les circonstances dans lesquelles elle a été forgée, d'attester les sources, les usages de référence. Les questions ou remarques soulevées par la définition - soit au sein de la recherche en EIAH, soit dans les disciplines associées - sont documentées dans la mesure où il est possible de les retrouver sous la forme de publications accessibles. Enfin, les traductions dans diverses langues sont indiquées accompagnées de commentaires éventuels. Des termes associés sont proposés en lien avec la construction du thésaurus. Un exemple d'entrée type du dictionnaire est donné en Annexe 1.

Wikimedia (<http://commons.wikimedia.org/>) s'est imposé de façon naturelle pour construire le dictionnaire, en raison de la facilité d'édition collaborative exigeant peu de compétences informatiques de la part des auteurs et fournissant de bons moyens de traçabilité des évolutions des textes. Cependant, la politique éditoriale adoptée ne se conforme pas totalement à la philosophie wiki. Elle exploite la possibilité de définir des niveaux de droits d'écriture ne permettant pas l'intervention de tout

internaute, la consultation étant bien sûr ouverte. Un effet de bord de ce choix est la restriction du droit d'accès à l'espace de discussion pour le visiteur de base, mais elle peut être compensée par la création d'un forum associé : (<http://www.linkedin.com/groups/TEL-dictionary-initiative-3880196>).

5.2. Le choix des mots

Choisir les entrées du dictionnaire est le résultat d'un processus empirique, dont le point de départ est la liste des mots clés associés par les auteurs au moment du dépôt de leurs publications dans l'archive TeLearn, complétée par le vocabulaire effectivement utilisé dans les textes. Nous avons constaté que les mots clés ne rendent pas compte aussi précisément qu'on pourrait l'imaginer du contenu effectif des textes ; l'analyse lexicale a révélé un ensemble de termes significatifs plus large. Comme cela est très généralement le cas, les fréquences sont très variables ; celles des termes potentiellement les plus significatifs pour le domaine sont faibles. Il s'agit donc, lors du choix, de prendre en compte ce qui est explicité par les auteurs, ce que l'analyse lexicale fait remonter à un certain seuil de fréquence, ainsi que des termes plus rares mais scientifiquement forts. Aussi l'approche a-t-elle été pragmatique : rassembler les termes à la fois les plus spécifiques et les plus fréquents.

Pour la première version du dictionnaire, preuve de concept, une liste de 100 termes a été retenue (Annexe 2). Le dictionnaire se déploiera ensuite pour inclure les entrées pertinentes qu'elles soient spécifiques des EIAH ou qu'elles correspondent à des concepts importants venant des domaines associés.

5.3. Structure éditoriale

Afin d'assurer un niveau de qualité aussi élevé que possible, une structure éditoriale a été constituée pour accompagner l'ensemble du projet et ses principales composantes. Le schéma présenté en figure 3 résume cette organisation.

Un conseil scientifique réunissant 19 chercheurs de divers horizons disciplinaires et culturels¹⁰ a été mis en place, attestant essentiellement

¹⁰ Rosy Bottino (IT), Paul Brna (UK), Tak-Wai Chan (TW), Charles Crook (UK), Ton de Jong (NL), Antonio Dias de Figueiredo (PT), Frank Fischer (DE), Monique Grandbastien (FR), Yasmin Kafai (USA), Judy Kay (AU), Stefanie Lindstaedt (AT), Marcia Linn (USA), Chee Kit Looi (SG),

l'intérêt pour l'initiative et, pour les plus actifs, apportant des idées pour les premiers développements. L'équipe éditoriale du dictionnaire est formée des auteurs des entrées ; elle évolue au fil de la construction. Il en est de même de l'équipe éditoriale multilingue qui réunit les représentants régionaux responsables des traductions des termes et, éventuellement, des entrées correspondantes.

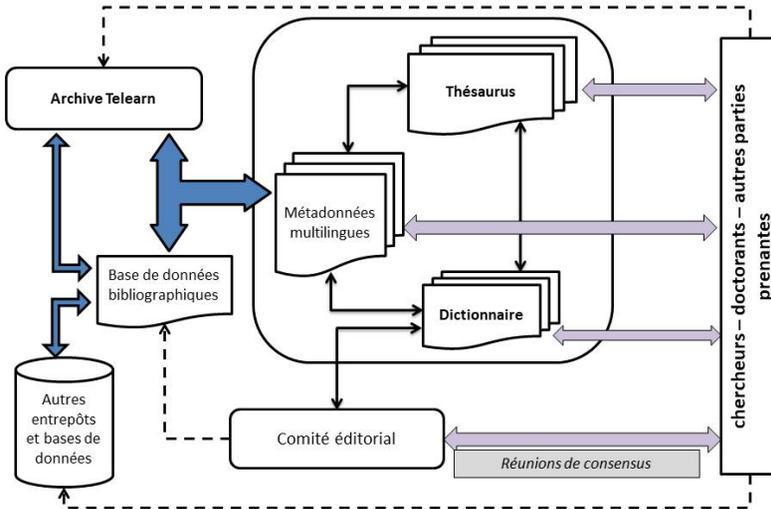


Figure 3 • Structure éditoriale du projet

6. Conclusion : un chantier nécessaire et ouvert

La pression sociale et économique, qui exige des résultats concrets et un impact tangible, est telle que la priorité de la recherche sur les EIAH est trop souvent donnée aux réalisations effectives au détriment du travail sur ce qui les fonde. Pourtant, ce sont ces fondements qui permettront le mieux de forger les moyens de concevoir et d'évaluer solidement les réalisations et, plus encore, de comprendre et de résoudre les problèmes qui réapparaissent de façon récurrente, quels que soient les progrès et innovations technologiques ou pédagogiques. Comprendre et modéliser les conditions d'un apprentissage efficace, caractériser opérationnelle-

ment les compétences et les significations, comprendre et décrire les conditions d'usage des EIAH au-delà d'une réalisation technique particulière, que la prochaine avancée technologique balayera sans pour autant résoudre ces problèmes fondamentaux, sont autant de questions pour lesquelles les réponses sont aujourd'hui trop partielles ou trop dispersées dans les divers champs disciplinaires pour pouvoir être mises en relation et intégrées. La constitution d'un cadre théorique ne signifie pas l'abandon de la pluralité, mais sa structuration par la mise en relation des concepts, des modèles et des théories associées souvent issus de diverses disciplines. Ce qui est en jeu est le statut de ces théories et méthodes dans le travail de la recherche et dans la formation des jeunes chercheurs. Ne pas saisir ce problème pour ce qu'il est conduit à une sorte de balkanisation en des écoles qui ont du mal à communiquer entre elles, et donc ont du mal à tirer le meilleur bénéfice des efforts réalisés sur ce qui, de fait, est un projet commun.

Travailler le vocabulaire, c'est-à-dire le sens des mots, leurs relations et leurs usages, est une première étape. L'enjeu est bien de travailler les concepts sous-jacents et ainsi d'avancer d'une façon pragmatique sur le terrain théorique. C'est aussi de faciliter et de consolider la circulation des idées et le partage des résultats de la recherche.

Cependant, travailler le vocabulaire c'est travailler la part explicite de la conceptualisation sous-jacente aux publications. L'analyse textuelle et ontologique montre qu'il ne faut pas, surtout dans un domaine jeune, en rester là. La construction d'une ontologie a été explorée en s'appuyant, d'une part, sur l'ontologie de l'éducation OMNIBUS et la méthode associée (Bourdeau, Mizoguchi, Hayashi, Psyche et Nkambou, 2007) et, d'autre part, sur la version initiale du vocabulaire anglophone. L'objectif était de relier les termes du vocabulaire et de proposer l'ébauche d'une modélisation du domaine « *Technology Enhanced Learning* » (Bourdeau et Balacheff, 2014). Selon la méthodologie OMNIBUS, la première tâche a été de cerner la signification de « *enhancement* » et de ses articulations avec « *technology* » et « *learning* ». La proposition, dans le cadre d'OMNIBUS, est de définir TEL comme un « *enhancing mechanism* » dont il reste à créer les catégories, à les relier et à les classer. La construction de cette ontologie en partant d'une autre dénomination du domaine pourrait avoir un autre résultat. Ce constat n'est pas singulier, il est fait pour d'autres domaines. Il signifie que, d'une part, le thésaurus, le dictionnaire et l'ontologie doivent être construits de façon collaborative et que, d'autre part, *leur fonction est*

descriptive et non normative (Aussenac-Gilles, Biébow et Szulman, 2000, section 5).

La mise en ligne du dictionnaire et du thésaurus, dont la réalisation actuelle est partielle, a suscité un intérêt dont attestent les scores des pages les plus populaires et les statistiques de fréquentation (1 375 905 pages depuis 2011¹¹). La fréquentation des pages traduites en chinois est particulièrement soutenue (chinois classique et simplifié) ainsi que celles traduites en français et en bulgare, pour lesquelles sont disponibles à la fois le vocabulaire et les entrées du dictionnaire.

Quoique le projet soit bien reçu, il est remarquable qu'il ne suscite jusqu'ici qu'un engagement limité des chercheurs. Une objection fréquente est l'existence de Wikipédia et, de fait, cette ressource est largement utilisée notamment par les doctorants. Cependant, la disparition de l'entrée « *e-learning* » et l'absence d'une entrée « *Technology Enhanced Learning* », expressions redirigées vers « *Educational technology* », montre de façon suffisamment évidente les limites de cette ressource. Dans ce domaine pas encore solidement fondé au plan scientifique, dans lequel l'intime conviction peut parfois se substituer à la preuve rationnelle, le contenu et l'organisation d'une encyclopédie sur le modèle Wikipédia peut souffrir de jeux d'influence au détriment de la qualité scientifique. La création d'une ressource ouverte, mais modérée, est nécessaire, c'est une responsabilité importante de la communauté de recherche. Le chantier est ouvert, déjà bien engagé, le mener à bien est moins une question de moyens que celle d'un engagement collectif. L'accueil du projet sur la plateforme de l'ATIEF et les effets de synergie que peut susciter l'association sont, dans cette perspective, très favorables.

REMERCIEMENTS

La conception et les premières réalisations de l'archive ouverte Telearn, du dictionnaire et du thésaurus ont été financées par le Réseau d'excellence Kaleidoscope (Communauté européenne, 6^e PCRD-IST, Projet N° 507838, 2004-2007), puis par le Réseau d'excellence STELLAR

¹¹ Vérifié à 10:21 le 19 octobre 2018.

(European Community, 7^e PCRD-ICT, Projet N° 231913, 2009-2012). Équipe technique: Sophie Min puis Émilie Manon pour l'expertise documentaliste, Jérôme Zeiliger pour la réalisation informatique. Une première présentation du projet a été publiée dans *TD Technologie Didattiche* (Balacheff *et al.*, 2012).

RÉFÉRENCES

Anthony, L. (2004). AntConc: A learner and classroom friendly, multi-platform corpus analysis toolkit. Dans L. Anthony, F. Shinichi et H. Yasunari (dir.), *Proceedings of IWLeL 2004: An Interactive Workshop on Language e-Learning* (p. 7-13). Tokyo, Japon : Waseda University. Récupéré de https://www.researchgate.net/profile/Laurence_Anthony/publication/267631346_Proceedings_of_IWLeL_2004_An_Interactive_Workshop_on_Language_E-learning_2004/links/5458cd870cf26d5090acf212.pdf#page=7

Anthony, L. (2010). AntConc, A freeware corpus analysis toolkit for concordancing and text analysis [logiciel]. Récupéré du site de l'auteur : <http://www.laurenceanthony.net/software/antconc/>

Artigue, M., Cerulli, M., Haspekian, M. et Maracci, M. (2009). Connecting and Integrating Theoretical Frames: The TELMA Contribution. *International Journal of Computers for Mathematical Learning*, 14(3), 217-240.

Aussenac-Gilles, N., Biébow, B. et Szulman, S. (2000). Modélisation du domaine par une méthode fondée sur l'analyse de corpus. Dans P. Tchounikine (dir.), *Actes de la 9e Conférence Francophone d'Ingénierie des Connaissances (IC 2000)* (p. 93-104). Toulouse, France : Université Paul Sabatier. <hal-00510453>

Balacheff, N., Baron, M., Desmoulins, C., Grandbastien M. et Vivet, M. (1997). Conception d'environnements informatiques d'apprentissage avec ordinateur. Dans S. Pesty, P. Siegel (dir.), *Actes des 6^e journées nationales PRC-GDR Intelligence artificielle* (p. 315-337). Paris, France : Hermès.

Balacheff, N. et Lund, K. (2012). Multidisciplinary vs multivocality, the case of "learning analytics". Dans *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (p. 5-13). New-York, NY : ACM. [hal-00805362]

Balacheff, N., Manon, E. et Zeiliger, J. (2012). Per un dizionario della ricerca sul Technology Enhanced Learning. *TD Technologie Didattiche*, 20(3), 155-161. [hal-00986096]

Baron, M. (1994). EIAO, quelques repères. *Terminal*, 65. Récupéré de http://www.revue-terminal.org/www/terminal-archives/articles/65/multimedia_baron.html

Baron, M., Gras, R. et Nicaud, J.-F. (1991). Introduction. Dans *Actes des Deuxièmes Journées EIAO* (p. 7-8). Cachan, France : Éditions de l'École Normale Supérieure de Cachan.

Bourdeau, J. et Balacheff, N. (2014). Technology-Enhanced Learning: From thesaurus and dictionary to ontology. Dans J. Jovanović et R. Chi-ong. (dir.),

Technological and Social Environments for Interactive Learning (p. 1-33). Santa Rosa, CA : Informing Science. Récupéré de <http://r-libre.telugu.ca/785/>

Bourdeau, J. et Grandbastien, M. (2011). La modélisation du tutorat dans les systèmes tutoriels intelligents. *Sciences et technologies de l'information et de la communication pour l'éducation et la formation*, 18. Récupéré de http://sticf.univ-lemans.fr/num/vol2011/02-bourdeau-tice/sticf_2011_tice_bourdeau_02.htm

Bourdeau, J., Mizoguchi, R., Hayashi, Y., Psyche, V. et Nkambou, R. (2007). When the domain of the ontology is education. Dans *Proceedings of the 4th Conference on Intelligent, Interactive Learning Objects Repository Networks (I2LOR'07)*, Montréal. Récupéré de http://www.lornet.ca/Portals/10/I2lor07_Proceedings/Tuesday/Papers/4-2.pdf

Bruillard, E. (1997). *Les machines à enseigner*. Paris, France : Hermès.

Carr, B. et Goldstein, I. P. (1977). *Overlays: a theory of modelling for computer aided instruction*. AI Memo 406. Cambridge, MA : Massachusetts Institute of Technology. Récupéré de <http://dspace.mit.edu/bitstream/handle/1721.1/5772/AIM-406.pdf?Sequence=2>

Collectif. (2003). *Concepts and methods for exploring the future of learning with digital technologies*. Kaleidoscope, initial proposal part B. European Commission, 6th Framework Program Network of excellence (IST-507838—2003-2007).

Desmoulins, C., Marquet, P. et Bouhineau, D. (2003). Introduction. Dans *Actes de la conférence Environnements informatiques pour l'apprentissage humain (EIAH 2003)*. Paris, France : INRP. Récupéré de <http://archiveseiah.univ-lemans.fr/EIAH2003/>

European Commission (2002). Report on the evaluation of expression of interest in the research priority 1.1.2i "Societal and economic challenges", sub-priority "Research addressing work and business challenges", Domain: Technology Enhanced Learning (pp. 31-39). Dans European Commission DG IST (dir.), *A European Research Area for Technology Enhanced Learning*. Information day 5 December 2002, Luxembourg.

Hlynka, D. (2013). Educational technology. Dans C. R. Richey (dir.), *Encyclopedia of terminology for educational communications and technology* (p. 102-104). New York, NY : Springer.

Journal officiel (2005, 14 mai). Formation en ligne. Récupéré du site : <http://www.culture.fr/franceterme/result?francetermeSearchTerme=elearning&francetermeSearchDomaine=0&francetermeSearchSubmit=rechercher&action=search>

Laborde, J.M. (2003). Des connaissances abstraites aux réalités artificielles, le concept de micromonde Cabri. Dans *Actes de la conférence Environnements Interactifs d'Apprentissage avec Ordinateur* (tome 2, p. 29-41). Paris, France : Eyrolles.

Lafontaine, J. (2011). Quelques mots de l'algèbre. Récupéré du site *Images des mathématiques*: <http://images.math.cnrs.fr/Quelques-mots-de-l-algebre.html>

Lajoie, S. P. et Naismith, L. (2011). Computer-based learning. Dans N. M. Seel (dir.), *Encyclopedia of the science of learning* (p. 716-718). Berlin, Allemagne : Springer.

Lowenthal, P. et Wilson, B. G. (2010). Labels DO Matter! A critique of AECT's redefinition of the field. *TechTrends*, 54(1), 38-46. <https://doi.org/10.1007/s11528-009-0362-y>

Manson P. (2007). Technology Enhanced Learning: Supporting learning in the 21st century. *ERCIM News - Special Technology-Enhanced Learning*, 71, 3 (Keynote). Récupéré de <https://ercim-news.ercim.eu/images/stories/EN71/EN71-web.pdf>

Min, S., Balacheff, N. et Zeiliger, J. (2008). TeLearn, une archive ouverte multilingue dans le domaine des technologies pour l'apprentissage. *AMETIST: Appropriation, Mutualisation, Expérimentations des Technologies de l'IST*, INIST, 2008, 2. <hal-00962047>

Moore, J. L., Dickson-Deane, C. et Galyen, K. (2011). e-Learning, online learning, and distance learning environments: Are they the same? *Internet and Higher Education*. 14(2), 129-135.

Morin, E. (1991). *La méthode 4. Les idées. Leur habitat, leur vie, leurs mœurs, leur organisation*. Paris, France : Seuil.

Pavel, S. (2002). Lexique de l'apprentissage en ligne (bilingue Anglais-Français). Bulletin de terminologie, 251. Ottawa: Travaux publics et Services gouvernementaux Canada. Récupéré de <http://publications.gc.ca/pub?id=9.582458&sl=1>

Reiser, R. A. (2007). What field did you say you were in? Defining and naming our field. Dans R. A. Reiser et J. V. Dempsey (dir.), *Trends and issues in instructional design and technology* (2nd ed., pp. 2-9). Upper Saddle River, NJ : Pearson.

Sangrà, A., Vlachopoulos, D. et Cabrera, N. (2012). Building an inclusive definition of e-learning: An approach to the conceptual framework. *The International Review of Research in Open and Distributed Learning (IRRODL)*, 13(2). <https://doi.org/10.19173/irrodl.v13i2.1161>

Vivet, M. et Nicaud, J.F. (1988). Les tuteurs intelligents : réalisations et tendances de recherches. *Technique et Sciences Informatiques*, 7(1), 21-45.

ANNEXE 1 – Un exemple d'entrée du dictionnaire

Fouille de données éducatives

Version 1

Édition : Michel C. Desmarais, Polytechnique Montréal

Contribution : Ryan S.J.d. Baker, Worcester Polytechnic Institute

Définition

La fouille de données éducatives est un processus conçu pour l'analyse de données issues de situations éducatives pour mieux comprendre les élèves et les situations dans lesquelles ils apprennent.

Jalons historiques

Cette expression est apparue dans le cadre d'une série de workshops sur le thème de l'analyse des usages (log data) par des élèves, dont le premier remonte à la conférence ITS 2000 à Montréal. En 2005, le premier workshop intitulé « Educational Data Mining » (EDM, Fouille de données éducatives) a été organisé à Pittsburgh conjointement à la conférence AAAI (Association for the Advancement of Artificial Intelligence). Depuis 2008, EDM désigne une conférence internationale organisée annuellement. Par ailleurs, une conférence sur la thématique voisine des Learning analytics (LAK2011) est apparue en 2011. En 2009 a été publié le premier numéro du journal international EDM, publication en accès direct disponible à l'URL <www.educationaldatamining.org/JEDM>. La société Educational Data Mining a été fondée en 2011 pour organiser la conférence et le journal.

Termes et expressions associés

Analytique des données éducationnelles, analyse des usages, données sur l'apprenant, donné d'usage de l'apprenant, découverte de la connaissance, fouille de données, apprentissage automatique statistique, psychométrie, modélisation de l'apprenant, classification, régression, classification, analyse factorielle, recherche de règles associatives, visualisation, découverte dirigée par les modèles, base de données.

Traduction

English: Educational data mining

Remarques disciplinaires

Le domaine de la fouille de données éducative (EDM) s'appuie sur des méthodes et théories issues de de nombreuses disciplines telles que la fouille de donnée, la découverte de connaissance, la psychométrie et l'apprentissage statistique. Il a pour but de construire des modèles et produire des résultats qui puissent aider à la conception et la réalisation d'applications et d'environnements innovants pour l'apprentissage, ainsi que d'apporter une contribution théorique à la psychologie de l'éducation ou d'autres domaines en éducation. Les méthodes pour l'EDM incluent (sans y être limité) la classification, la régression, l'analyse factorielle, la fouille de relation, la prédiction de connaissances, la fouille de corrélations, la recherche de règles associatives, la visualisation, la découverte de structure de domaine, la découverte dirigée par des modèles.

Principales références

- Baker R., Yacef K. (2009) The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, 1, 3-17.
- Baker R.S.J.d. (2010) Data Mining for Education. In: Mcgaw B, Peterson P., Baker R. (eds.) *International Encyclopedia of Education* (3rd edition) (7, 112-118). Oxford, UK: Elsevier.
- Romero C., Ventura, S. (2007) *Educational Data Mining: A Survey from 1995 to 2005*. *Expert Systems with Applications*, 33, 125-146.
- Romero C., Ventura S. (2010) *Educational Data Mining: A Review of the State-of-the-Art*. *IEEE Transaction on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 6, 601 - 618.
- Koedinger K.R., Cunningham K. A. S., Leber B. (2008) An open repository and analysis tools for fine-grained, longitudinal learner data. In: *Proceedings of the 1st International Conference on Educational Data Mining*, 157-166.
- EDM, International Working Group on Educational Data Mining: Educational data mining. <http://www.educationaldatamining.org/> (2009)

ANNEXE 2 – Vocabulaire initial

<p>A</p> <ol style="list-style-type: none">1. Adaptive learning environment2. Agent-based tutoring system3. Agent-based learning environment4. Ambient learning environment5. Animated pedagogical agent6. Artificial learning environment7. Augmented learning environment8. Authoring system9. Authoring tool10. Automatic diagnosis11. Automatic generation of exercises <p>B</p> <ol style="list-style-type: none">12. Blended learning13. Buddy systems <p>C</p> <ol style="list-style-type: none">14. Cognitive diagnosis15. Cognitive engineering16. Cognitive modeling17. Cognitive tutor18. Collaborative learning19. Computational mathetic20. Computational thinking21. Computer-assisted examination22. Computer-assisted instruction23. Computer-assisted language learning24. Computer-assisted learning25. Computer-based instruction26. Computer-based learning environments27. Computer-based laboratory28. Computer programming in support of learning29. Computer-supported collaborative learning30. Constructionism31. Course management system32. Courseware	<p>D</p> <ol style="list-style-type: none">33. Didactical engineering34. Distance learning35. Distributed learning36. Distributed learning environments37. Dynamic geometry <p>E</p> <ol style="list-style-type: none">38. e-Assessment39. Educational affordance40. Educational data mining41. Edutainment42. e-learning43. Embedded phenomena44. Epistemic affordance45. Epistemic feedback46. e-portfolio47. Experiential learning <p>G</p> <ol style="list-style-type: none">48. Game-based learning <p>I</p> <ol style="list-style-type: none">49. Inclusive e-learning50. Informal learning51. Inquiry learning52. Intelligent scaffolding system53. Intelligent tutoring system54. Interactive learning environment55. Interactive white board56. Integrated learning system <p>L</p> <ol style="list-style-type: none">57. Learner modeling58. Learning analytics59. Learning companion60. Learning design61. Learning environment62. Learning grid63. Learning management systems64. Learning objects65. Learning scenario66. Learning space67. Learning trail <p>M</p> <ol style="list-style-type: none">68. Microworld69. Mobile learning70. Multimedia learning	<p>N</p> <ol style="list-style-type: none">71. Narrative learning environments72. Networked learning73. --- communities74. --- environment <p>O</p> <ol style="list-style-type: none">75. On-line learning environment76. Overlay <p>P</p> <ol style="list-style-type: none">77. Pedagogical agent78. Personal learning environment79. Pervasive learning environment80. Programmable computer-based learning environment81. Programmable course <p>R</p> <ol style="list-style-type: none">82. Remote laboratories <p>S</p> <ol style="list-style-type: none">83. Script ...84. --- collaboration85. --- external86. --- internal87. Seamless learning environment88. Self-regulated learning89. Simulation games90. Simulation-based learning environment91. Situated learning <p>T</p> <ol style="list-style-type: none">92. Tangible learning environment93. Technology Enhanced Learning94. Techno-mathematical literacies <p>U</p> <ol style="list-style-type: none">95. Ubiquitous learning <p>V</p> <ol style="list-style-type: none">96. Virtual campus97. Virtual learning environment98. Virtual pedagogical agent99. Virtual universities <p>W</p> <ol style="list-style-type: none">100. Web-lecturing technologies
--	---	--

