HAL open science

TeLearn, une archive ouverte multilingue dans le domaine des technologies pour l'apprentissage

S. Min, Nicolas Balacheff, Jérôme Zeiliger

To cite this version:

S. Min, Nicolas Balacheff, Jérôme Zeiliger. TeLearn, une archive ouverte multilingue dans le domaine des technologies pour l'apprentissage. AMETIST : Appropriation, Mutialisation, Expérimentations des Technologies de l'IST, 2008, 2, pp.n.a. hal-00962047

HAL Id: hal-00962047

https://hal.science/hal-00962047v1

Submitted on 5 Jan 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

version auteur

TeLearn, une archive ouverte multilingue dans le domaine des technologies pour l'apprentissage

Sophie MIN, Nicolas BALACHEFF, Jérôme ZEILIGER

Laboratoire LIG - 46, Av F. Viallet 38031 Grenoble Cedex

Sophie.min@imag.fr

Résumé : TeLearn est une archive ouverte cherchant à rassembler des communautés de chercheurs qui travaillent sur l'apprentissage assisté par les technologies ou en anglais « Technology Enhanced Learning » - TEL. Le but est de permettre un accès unique et multilingue à des ressources scientifiques hétérogènes (textes, vidéos, images, données) et la conception de services répondant aux attentes professionnelles spécifiques à leur secteur scientifique.

Cet article présente le projet TeLearn en termes de fonctionnement, de statistiques sur les données et les utilisateurs et les perspectives de développement.

Mots-clés : archive ouverte, système, technologies pour l'apprentissage

---

Abstract: TeLearn is an open archive developed for research on the field of "Technology Enhanced Learning "- TEL. And it is a unique, openly, freely, and reliably access to TEL research. The TeLearn archive accepts material in any language, and the resources found here include not only publications, but also audio and video presentations. Moreover services have been designed to own research expectations.

This article explains how TeLearn works and statistical data of users and resources. Afterwards, we present development future prospects.

Keywords: open archive, system, technology enhanced learning

1 Présentation du projet

1.1 Kaléidoscope et TeLearn

TeLearn est né, en 2006, de la volonté du réseau européen de recherche Kaléidoscope[1] de « permettre une communication la plus ouverte possible des résultats de la recherche » [1]. En effet, le but d'un réseau d'excellence est de renforcer et de structurer la recherche européenne ; mais aussi contribuer à la capitalisation et au transfert de connaissance. Kaléidoscope rassemble des laboratoires dans le domaine des technologies de l'éducation et de la formation. Cette thématique est à l'intersection de plusieurs disciplines dont notamment les sciences de l'éducation, les sciences de l'homme et de la société et les sciences et technologies de l'information et de la communication (STIC).

Cette pluridisciplinarité et le multilinguisme font de TeLearn une archive ouverte originale.

1.2 Objectifs

Le principe des archives ouvertes est simple. Il s'agit de mettre à disposition le plus largement et rapidement possible les résultats de la recherche scientifique sans contraintes d'accès. On doit aux physiciens, à la fin des années 90, l'initiative de cette politique de partage des connaissances scientifiques en construction. Les archives ouvertes ont été adoptées internationalement par plusieurs disciplines, outre la discipline fondatrice, la physique, on peut mentionner les mathématiques, la biologie et l'informatique. D'autres disciplines, telles les sciences cognitives, ont adopté la même stratégie, mais il revient à la France d'avoir ouvert la voie des archives ouvertes à toutes les disciplines, dont celles des SHS. Tous les secteurs disciplinaires ne réagissent pas de la même façon à la proposition d'adopter ce moyen de diffusion de leurs travaux, malgré les incitations institutionnelles et par exemple l'appel de Berlin au

[1] Kaléidoscope est un réseau d'excellence européen créé en janvier 2004 dans le cadre du sixième programme cadre. Il rassemble près d'une centaine de laboratoires représentant plus d'un millier de chercheurs dans le domaine des TEL.

début des années 2000. Les chercheurs du secteur SHS, notamment, paraissent avoir de fortes réticences [2] Toutefois on remarque des initiatives significatives dans le domaine particulier des recherches sur les technologies pour l'apprentissage, l'éducation et la formation. Aussi était-il naturel de relever le défi de monter une archive ouverte dans ce domaine, au niveau international où rien n'était fait, au sein du réseau d'excellence Kaleidoscope piloté par le CNRS dans le contexte du 6ème programme cadre de la communauté européenne.

C'est ainsi qu'est née, en 2005, l'idée de créer TeLearn, archive ouverte internationale, multilingue et multidisciplinaire (des STIC aux SHS) pour la diffusion des résultats des recherches sur la conception, la réalisation, l'usage et le déploiement des technologies d'éducation et de formation.

La suite de ce texte présente les grandes étapes de cette construction, l'architecture de TeLearn, des éléments sur son usage et finalement les perspectives alors que le projet donne des signes d'une adoption significative et a par ailleurs fait l'objet d'une évaluation très favorable de la part de la commission européenne lors de l'examen final du réseau Kaleidoscope.

1.3 Les débuts

En 2005, le projet du réseau Kaleidoscope était de proposer des services autour d'un corpus de documents permettant un partage et une communication entre chercheurs. La constitution d'une archive ouverte s'est imposée comme la réponse la plus naturelle à cet objectif.

Avant de mettre en place une plateforme « grand » public, il est apparu judicieux de récupérer quelques documents à l'intérieur

du réseau et de travailler sur une liste commune et pertinente de mots-clés.

Une première version « brouillon » a été mise en place utilisant la plateforme technique et le savoir faire du CCSD - Centre pour la Communication Scientifique Direct du CNRS, en termes de métadonnées. Cette instance a servi de première bouture au système actuel de TeLearn. Même si le savoir faire du CCSD a permis de formaliser les métadonnées utiles pour une archive ouverte, la rigidité technique ne permettait pas de concevoir des services à plus forte valeur ajoutée.

Le système actuel de TeLearn a été développé sur les ressources propres du réseau Kaleidoscope pour s'affranchir des contraintes techniques du CCSD ; avec cependant la particularité de maintenir un entrepôt miroir à HAL[2] pour assurer une conservation pérenne des ressources. C'est pourquoi une base de données et une plateforme ont été créées.

1.4 TeLearn dans le panorama des archives ouvertes

TeLearn s'inscrit naturellement dans l'esprit de l'« Open Archives Initiative » (OAI) promouvant l'échange et la communication libre et gratuit des produits de la recherche. C'est pourquoi la structure d'une fiche de métadonnées de TeLearn est décrite en Dublin Core et est conforme aux recommandations du protocole Open Archive Initiative Protocol for Metadata Harvesting (OAI-PMH).

Dans le panorama des archives ouvertes internationales existantes[3], TeLearn parait être le seul cas d'entrepôt de ressources multilingues et pluridisciplinaires sur les domaines des technologies de l'apprentissage. Il existe bien des ressources travaillant sur des

[2] Hyper Article en Ligne.

[3] D'après le répertoire DMOZ et la liste de www.openarchives.org

thématiques proches comme l'archive Edutice[4] ou l'archive Redined[5] mais elles sont spécifiques de la recherche française ou espagnole.

2 La conception de TeLearn

2.1 Le fonctionnement

La dynamique du système est basée sur la réaction de différents acteurs, sa modération et des services. TeLearn est accessible à l'adresse : http://www.telearn.org[6]

2.1.1 Les acteurs

Le système TeLearn fonctionne grâce à l'interaction et la participation de cinq types d'acteurs différents. Cette typologie est basée selon les droits d'accès et d'actions possibles sur le système. Le visiteur, le déposant, l'administrateur (scientifique, informatique et de contenu); et de deux entités : l'institution du déposant et le CCSD qui assure la conservation des dépôts.

Le visiteur

Le visiteur est toute personne sans droits particuliers vis-à-vis du système. Il a accès aux services de base qui sont la consultation, le téléchargement et la recherche de ressources sur le site web TeLearn et auxquelles s'ajoutent des services complémentaires que nous décrirons plus loin.

Le déposant

Le déposant est la personne qui réalise l'action de remplir la fiche de métadonnées d'une publication. Cette personne doit obligatoirement être inscrite. En plus des services de bases,

[4] http://edutice.archives-ouvertes.fr/

[5] http://www.redined.mec.es/

[6] Plusieurs noms de domaine permettent d'accéder au système : telearn.noe-kaleidoscope.org .net - .com - .eu ; www.telearn.org - .net - .eu ; www.telearn-archive.org - .net - .eu ; telearn.net - .eu ; telearn-archive.org - .net - .eu

l’inscription permet de pouvoir modifier une ressource déposée et de suivre l’état de ses dépôts.

Le déposant peut être différent de l'auteur d'une publication. Un auteur TeLearn est une personne ayant donné son accord pour diffuser ses publications à travers l’archive. Ce fonctionnement permet aux auteurs de déléguer le travail de dépôt. L’archive TeLearn ne gère pas les droits d’auteurs ; le dépôt d’une ressource est placé sous la responsabilité de la personne qui l’effectue ; il implique l’accord de l’ensemble des auteurs. Ces derniers conservent évidemment l’entière propriété intellectuelle de leur travail.

L’institution du déposant

Pour qu’un déposant puisse s’inscrire, il faut que son institution ou unité de recherche soit affiliée à TeLearn. Un représentant du responsable de l’unité de recherche remplit une fiche permettant de valider l’adéquation des thèmes de recherche à ceux traités par TeLearn.

Le responsable de l’institution des déposants a des droits d’accès à la fiche de son institution et de création/modification des tampons associés. Un tampon est une marque apposée sur les ressources pour les identifier comme appartenant à l’institution ou pour définir une valeur scientifique.

L’administrateur

Trois types d’administrateurs assurent le bon fonctionnement du projet d’un point de vue scientifique, informatique et du contenu. Leurs droits d’accès et d’actions sur le système ne sont, par conséquent, pas limités.

Le manager scientifique définit le périmètre scientifique des ressources et la politique globale d’évolution de l’archive. L’ingénieur informaticien assure le développement du système. Le gestionnaire de contenu assure la cohérence des données et aide les utilisateurs sur les aspects techniques et sur les questions relatives aux droits d’auteur. C’est aussi l’animateur du site Web. Son rôle,

essentiellement lors de la phase de lancement, est d'inciter les dépôts par une communication avec les chercheurs.

Le CCSD

TeLearn travaille en partenariat avec le CCSD pour garantir la pérennisation des données du domaine de HAL : Informatique ; sous domaine : *Environnements Informatiques pour l'Apprentissage Humain* (EIAH). Techniquement, l'importation des données TeLearn vers HAL est réalisée grâce aux webservices du CCSD et à l'importation XML.

2.1.2 La soumission

La soumission est le processus réalisé par une personne enregistrée pour solliciter le dépôt d'une ressource dans TeLearn. Ces ressources sont pour l'instant de deux types : publication et vidéo.

Le formulaire de soumission est découpé en plusieurs sections selon les types de ressources, permettant de structurer les dépôts relativement aux métadonnées à saisir.

La zone de soumission qui permet de trouver le formulaire récapitule aussi l'activité du déposant en lui listant les différents états des ressources déposées.

L'étape de la soumission est l'une des premières actions réalisables par le déposant. Le diagramme suivant (Figure 1) présente les différentes actions possibles des acteurs qui influent sur l'état d'une ressource.

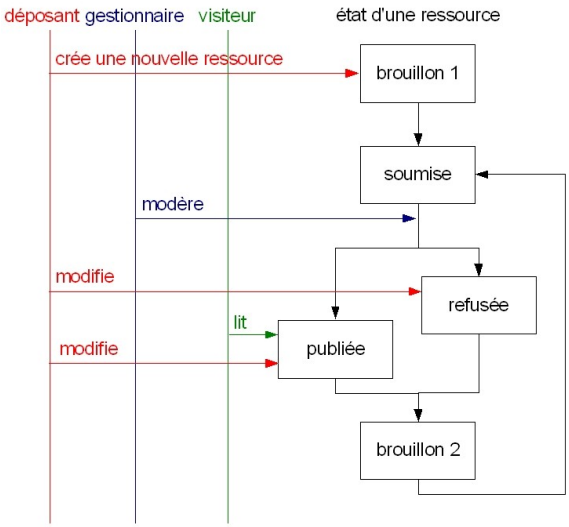


Figure 1 : Diagramme des actions possibles des acteurs pendant la vie d'une ressource

2.1.3 La modération

Pour maximiser la pertinence, l'archive TeLearn est modérée à deux niveaux :

- les déposants;
- les ressources (publication ou vidéo).

Le but de la modération est de limiter le bruit sans se substituer aux experts scientifiques du domaine. La valeur scientifique du contenu des ressources n'est donc pas évaluée comme pourrait le faire un comité de lecture.

2.1.3.1 Modération des institutions déposantes

La modération des unités de recherche permet de faire une présélection des ressources déposées ; en partant du principe que les déposants des unités de recherche validées entreposeront des

ressources pertinentes pour TeLearn. Le critère de validation d'une candidature à l'affiliation est que l'institution ait des activités de recherche, ou R&D dans le domaine d'intérêt de TeLearn.

L'inscription à TeLearn permet au référent du laboratoire d'identifier les publications de son unité de recherche ou de mettre en valeur un point de vue ; grâce à un système de tampons. Il désigne aussi quelqu'un pour « tamponner » des ressources. Les tampons peuvent être visibles ou invisibles.

2.1.3.2 Modération des ressources

La modération des publications est un second contrôle permettant de vérifier la pertinence de chaque publication soumise. En effet, les ressources déposées doivent avoir un contenu susceptible d'intéresser la communauté des chercheurs sur les technologies d'éducation et de formation.

La figure 2 formalise la procédure de validation d'une ressource mise en œuvre par le gestionnaire de contenu. C'est-à-dire les conditions de publication ou de refus d'une ressource soumise. Un refus est toujours motivé. Ces sept étapes permettent d'examiner quelques points clés, dans le but d'assurer une certaine qualité aux données. Les voici :

- le doublon possible dans la base ;
- la recevabilité scientifique ;
- la validité des fichiers (PDF ou vidéo) et des liens attachés ;
- les auteurs et leurs affiliations : nom du laboratoire au moment de la création de la ressource, l'ordre des auteurs ;
- les mots-clés.

Dans le cas des publications, l'archive TeLearn accepte les références bibliographiques, les liens vers d'autres sites web ou les textes intégraux sous format PDF. Cependant, le dépôt en texte intégral est recommandé. Dans le cas des vidéos, les données

peuvent être hébergées par le serveur de TeLearn ou sur un autre site. Les vidéos sont de préférence en streaming[7].

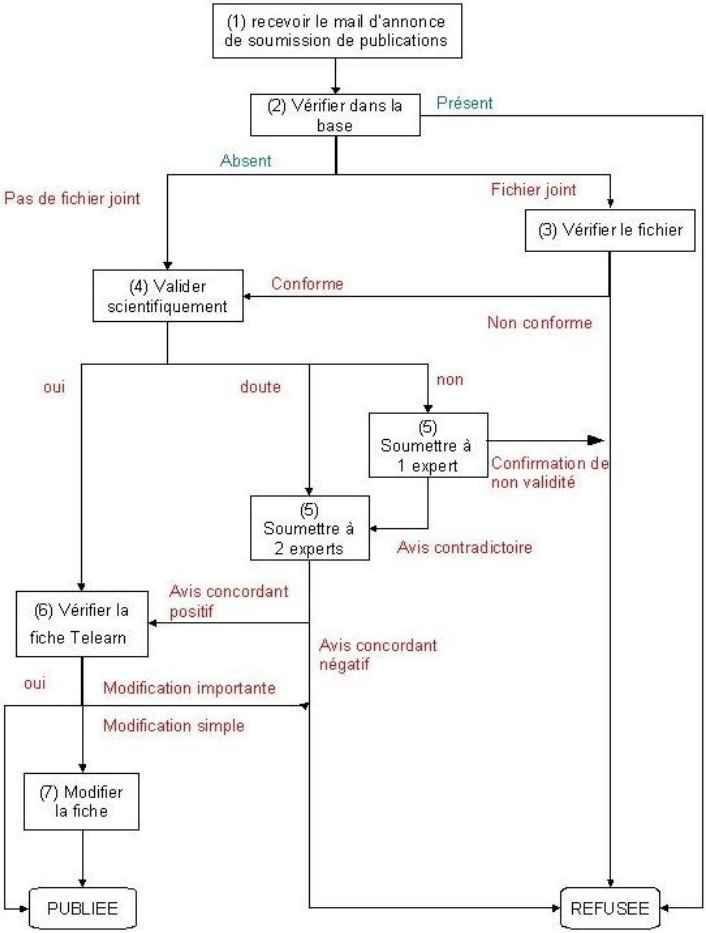


Figure 2 : Procédure de validation d'un document

[7] Le streaming permet de diffuser du son ou des vidéos en continu, ce qui implique un temps d'attente réduit.

2.1.4 Les autres services

Autour de l'archive TeLearn, il existe des services représentés en rouge sur la figure 3. En vert sont indiquées les différentes zones de la page d'accueil.

Figure 3 : Les autres services et la page d'accueil

- La **consultation** est segmentée par des filtres qui pré-trient les ressources. Par exemple, une consultation par auteurs est possible.
- La **recherche** est pour l'instant simple puisque qu'elle est effectuée sans opérateur booléen sur le titre, le résumé, les références et les mots-clés.
- L'**aide** propose une rubrique Questions-Réponses sur l'archive, les droits d'auteurs et des tutoriels sur la soumission et l'identification.

- L'**actualité** classée selon 3 thèmes : autour de TeLearn, le mouvement d'archive ouvert et la documentation dans le domaine.
- L'**annuaire de liens** présente les revues, de préférence en libre accès, intéressantes pour les chercheurs du domaine, classées par thème.
- Les **fils RSS** (Really Simple Syndication) permettent de suivre en temps réel l'activité de dépôt des différents types de ressources. Un fil RSS est aussi rattaché à chaque collection.

Pour permettre une diffusion fiable des URLs, elles sont statiques. Par exemple, pour une publication l'adresse est du type où [num] est le numéro de la publication dans la base et [ver] le numéro de la version :

http://www.telearn.org/open-archive/browse?resource=[num]_v[ver]

2.2 L'architecture technique

Le site www.telearn.org est géré via un système de publication web et de gestion de contenu (Content Management System), le CMS Alpha Complex également développé pour le projet Kaleidoscope.

Un site développé sous Alpha Complex est considéré comme un arbre dont chaque embranchement est un nœud. A chaque nœud est attaché un comportement (type page, type menu…), des paramètres (titre, auteur…), des niveaux d'accès (lecteur, contributeur, éditeur…) et un lien éventuel vers une base de données.

Alpha Complex est écrit en PHP objet et s'appuie sur une base de données MySQL. Son architecture est modulaire, c'est-à-dire que ses fonctionnalités génériques sont encapsulées sous forme de 'modules' qui peuvent être utilisés à volonté en des endroits divers des sites à développer (tels les modules 'Menu' ou 'News' ou 'Blog' ou 'File_depot' etc.).

Ainsi la fonctionnalité ‘archive ouverte’ est réalisée par le module ‘open archive’ présent à la racine du site et qui prend en charge toutes les fonctions de :

- parcours, tri et affichage du contenu de l’archive (‘browse’) ;
- recherche (‘search’) ;
- soumission et édition de ressources (‘submit’) ;
- modération (interface d’administration) ;
- gestion des institutions affiliées et des collections ;
- et l’interfaçage OAI.

Le système TeLearn est en effet un entrepôt de données conforme à la norme OAI-PMH, dûment référencé comme tel, ce qui le rend moissonnable par d’autres archives fournisseurs de services (le point d’entrée de l’interface OAI est http://telearn.noe-kaleidoscope.org/open-archive/oai).

Cette modularité du système fait que d’autres systèmes comme TeLearn pourraient très facilement être créés pour des archives ouvertes d’autres domaines scientifiques.

3 Premières statistiques et retour d'usage (Novembre 2007)

Cette partie a pour but de présenter l’état d’avancement de TeLearn. Les éléments présentés se réfèrent à la période de mai 2006 (début de la mise en place) à fin octobre 2007.

Les statistiques présentant l'activité sur le site web (visites, téléchargements,...) proviennent de l'analyseur de log web **AWStats**[8] [2]. Sa mise en place n’a été effective qu'en décembre 2006.

[8]Version 1.857

3.1 Statistiques sur les ressources

Au 30 octobre 2007, il y avait **940 ressources** dont 804 publications, 118 vidéos et 18 outils logiciels. L'utilisation est en moyenne de **516 publications téléchargées** sous format PDF (demande de fichier par le navigateur), une quinzaine de vidéos et 5 outils logiciels[9] par semaine.

Malgré le caractère multilingue de l'archive, la figure 4 montre que la majorité des ressources sont en anglais.

Langage	%
English	81.2
French	16.0
Italian	1.2
Spanish	0.4
Dutch	0.4
Bulgarian	0.3
Danish	0.2
Portuguese	0.2

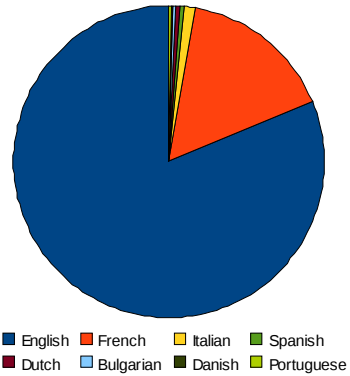


Figure 4 : Répartition des langues des ressources (en %)

Les ressources sont majoritairement des rétro-dépôt, des dépôts de ressources anciennes. La publication la plus ancienne date de 1991 et la vidéo est de 1997. L'année où les ressources sont le plus publiées est 2005. Il n'y a 117 ressources datant de 2007, soit **12,7% des dépôts** et 154 qui datent de 2006 (16,7%). Les années de publication ou de création d'un logiciel ne sont pas comptabilisées dans la figure 5. Une explication possible à ce constat est que les problèmes de copyright ne permettent pas aux auteurs de publier sans embargo. Ou les délais de publication d'une ressource sont longs par rapport à la date d'écriture ou d'acceptation.

[9] Dans le but d'intégrer toutes les ressources produites par la recherche, TeLearn a démarré cette initiative d'entrepôt des outils logiciels. Actuellement, le dépôt est possible mais pas encore mis en œuvre systématiquement.

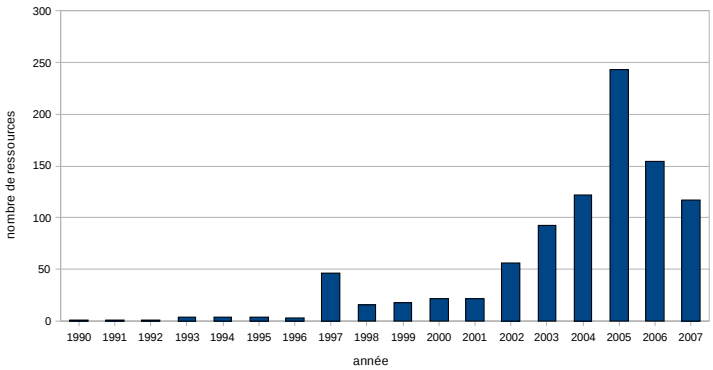
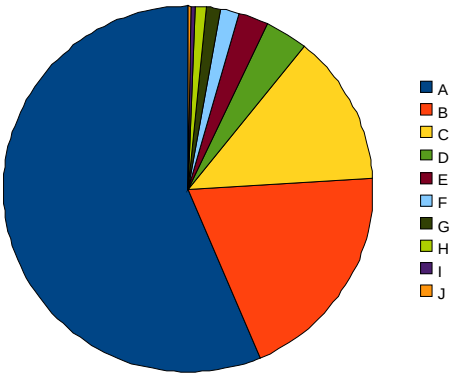


Figure 5 : Répartition des ressources publiées par date

Seulement 13,4% des dépôts sont des références bibliographiques. La plupart des publications déposées le sont en texte intégral hébergé par TeLearn (72,3%) ou sur un autre site (14,3%).

Les chercheurs déposent (voir figure 6) surtout des publications dans des conférences à comité de lecture (Publication/communication in peer-reviewed conference/proceedings) – **56,3%**, puis des publications dans des revues à comité de lecture – 19,7% (Publication in peer-reviewed journal) et enfin des rapports de recherche – 13,2%,. D'après la synthèse d'enquête d'Agnès Magron (Institut des Sciences de l'Homme – Lyon) [3] et en particulier l'enquête Alma Swan et Sheridan Brown [4] sur les différences de types de publication auto-archivées, les chercheurs en TEL se rapprocheraient plus de ceux en Informatique que ceux en Sciences sociales et de l'éducation où les publications circulent surtout en post-publication (donc article évalué par un éditeur).



Code	Publication type	%
A	Publication/communication in peer-reviewed conference/proceedings	56,3
B	Publication in peer-reviewed journal	19,7
C	Research report	13,2
D	Book chapter	3,7
E	Dissertation/thesis (PhD and Master level)	2,6
F	Seminar, workshop	1,6
G	Book	1,2
H	Other	1
I	Proceedings/transactions (whole papers of conference, workshop, congress or other meeting)	0,4
J	Publication in non peer-reviewed journal	0,2

Figure 6 : Quantité de publications par type de documents

Un des services mis à disposition pour les unités de recherche affiliées sont les tampons. Cette marque est aussi utilisée par le gestionnaire pour rassembler les différentes éditions d'une conférence. Actuellement, il a **19 tampons** dont 11 appartenant à des unités de recherche et 8 représentant des conférences. L'avantage d'un tampon est aussi de disposer d'une adresse stable pour accéder à une liste de ressources, mais seulement 3 groupes de recherche ont

utilisé ce lien. L'hypothèse pour répondre à cette sous-exploitation est que ce service n'est pas assez connu des utilisateurs.

3.2 Statistiques sur les utilisateurs

3.2.1 Les visiteurs

En moyenne sur la période de décembre 2006 à novembre 2007 (voir figure 7), il y a eu **1930 visiteurs différents** pour **4057 visites** par mois ; soit par semaine 483 visiteurs différents pour 1014 visites. En moyenne, un visiteur ne passe pas plus de 6 min sur TeLearn. Les deux derniers mois (septembre et octobre) montrent un nombre de visites et de visiteurs nettement significatif par rapport au reste de l'année.

Le nombre de visiteurs différents (unique visitors) représente le nombre d'ordinateurs depuis lesquels le site a été visité. Donc les personnes connectées par un réseau privé ne sont comptabilisées qu'une fois. Par contre, les visites (number of visits) précisent le nombre de fois où le site a été consulté, une visite est une nouvelle connexion au site. Les chargements de chaque page ne sont pas enregistrés ici.

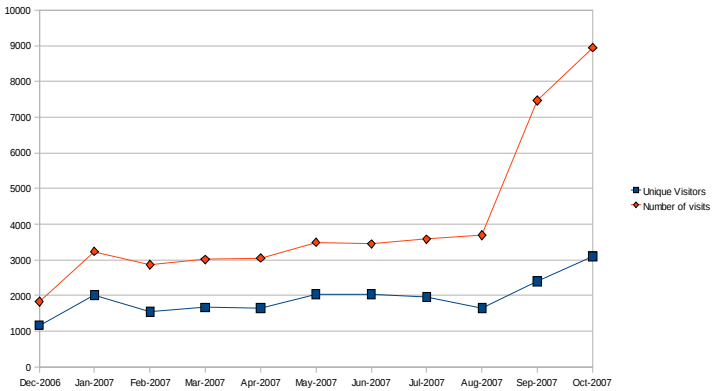


Figure 7 : Activité du site web TeLearn

Le diagramme suivant (Figure 8) présente les différentes origines de connexion des visiteurs aux pages du site. Quatre origines sont considérées :

- adresse directe ou favoris - Direct access / Bookmark ;
- Lien depuis un moteur de recherche Internet - Search engines ;
- Lien depuis une page externe (autres sites, hors moteurs) - Search engines ;
- Lien depuis une page externe (autres sites, hors moteurs) - externe webpage.

D'après les courbes de la figure 8, la très grande majorité des visiteurs sur le site se connecte en tapant ou cliquant sur une adresse du site présente dans ses favoris ou dans ses mails. En moyenne, huit fois moins de personne se sont connectés grâce à un moteur de recherche ; et vingt fois moins par une adresse externe. Ce qui laisse supposer que peu de sites pointe sur www.telearn.org. L'activité des connexions par adresse direct ou favoris et par moteurs de recherche s'est accentuée depuis le mois de juin.

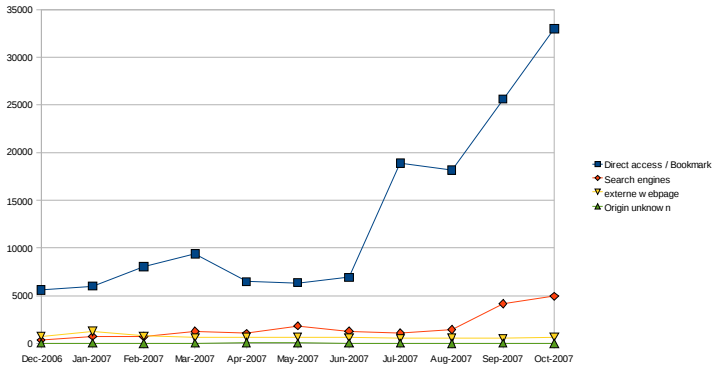


Figure 8 : origine des connexions sur www.telearn.org

3.2.2 Les laboratoires affiliés

Rappelons qu'un laboratoire ou unité de recherche affiliée est enregistré dans TeLearn et sont les seuls à pouvoir soumettre. Au 1er novembre, il y a **137 unités de recherche affiliées** dont 118 participent au réseau européen Kaleidoscope. Les membres de Kaleidoscope ont été affiliés par défaut. Cette répartition explique que la plupart des unités de recherche sont européennes à l'image du réseau d'excellence. (voir figure 9)

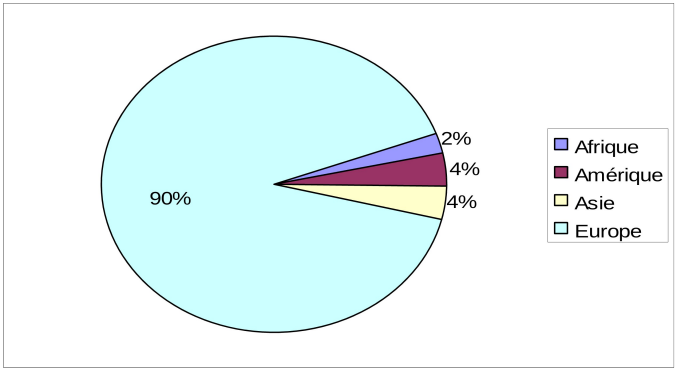


Figure 9 : Répartition des origines géographiques des laboratoires affiliés

Un peu plus de la moitié des laboratoires (75) participent activement au développement de l'archive. C'est-à-dire qu'au moins une personne de l'unité de recherche (secrétaire, documentaliste, thésard) a déposé une ressource.

3.2.3 Les déposants

Au 1er novembre, 214 personnes ont déposé au moins une ressource, soit en moyenne 4 ressources par déposant. Mais un peu moins de la moitié (47,2 %) n'ont publiée qu'une seule ressource. Les déposants qui ont publiés plus de 10 ressources (7,5%) sont le plus souvent des personnes à qui ont a délégué le travail.

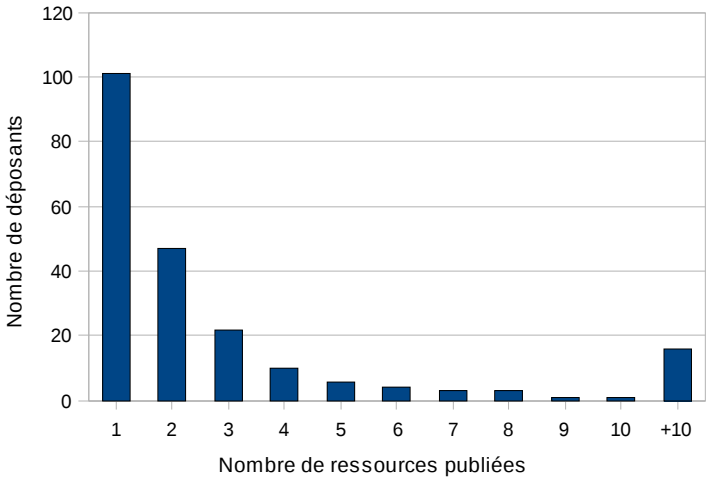


Figure 10 : Quantité de ressources publiées par déposants

Le diagramme suivant présente l'activité des déposants (Figure 11). En moyenne, il y a une cinquantaine de ressources qui sont déposées par mois. Une hausse significative des ressources (plus de 50% de la moyenne) à partir du mois d'août s'explique par la concrétisation de certaines actions de partenariats. En effet, au mois d'août, près de 80 publications provenaient d'un moissonnage du domaine de l'archive ouverte du CCSD, HAL : Informatique/Environnements Informatiques pour l'Apprentissage Humain. Au mois de septembre, le professeur Roy Pea a déposé près 80 publications représentant son travail sur l'apprentissage assisté par les technologies. Il y a eu aussi au mois d'octobre quelques accords pour déposer des vidéos de conférence s'étant déroulé pendant l'été.

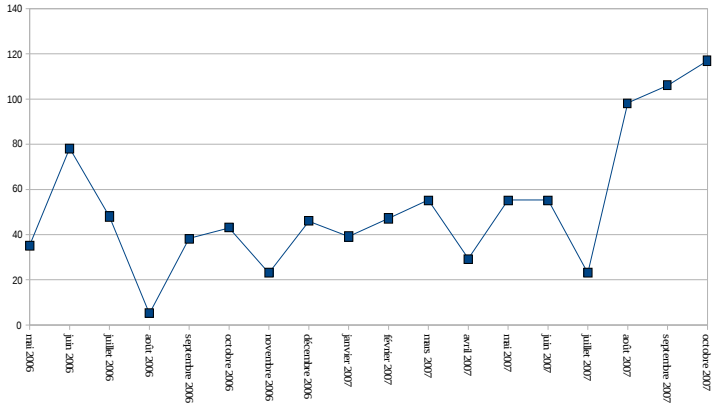


Figure 11 : Répartition du nombre de dépôts par mois

3.3 Portraits et retour d'usage

Les éléments statistiques ci-dessus permettent de dresser un portrait de l'archive et de ses utilisateurs. Cependant il faut le tempérer parce que moins de 2 ans d'existence ne rendent pas les chiffres significatifs sur l'appropriation réelle.

3.3.1 Portrait de l'archive et de ses utilisateurs

L'archive TeLearn est de taille respectable par rapport à son âge, un peu plus de 1 an et demi. Un visiteur du site télécharge au moins une publication sous format PDF, mais ne passe pas plus de 6 min sur le site.

Le déposant type appartient à une unité de recherche européenne qui appartient à la communauté Kaleidoscope. Son institution n'a pas de tampons. Il a 4 rétro-dépôt en anglais qui a six fois plus de chance d'être une publication qu'une vidéo ou un outil logiciel. S'il dépose une publication, c'est un papier d'une conférence à comité de lecture sous format PDF.

L'activité des derniers mois montre un début d'appropriation de l'archive parce que le nombre de visiteurs (et de visites) à augmenter significativement (plus de trois fois plus qu'en moyenne). Ces chiffres sont à rapprocher au nombre de dépôts deux fois plus important ces derniers mois, mais aussi aux connexions par adresse directe ou par favoris (vingt fois plus).

3.3.2 Retour d'usage informel

En juin 2007, un séminaire organisé sur TeLearn pour l'équipe Metah du Laboratoire d'Informatique de Grenoble (LIG) et membre du réseau Kaleidoscope a été l'occasion de connaître l'avis d'utilisateurs.

Les chercheurs ont montré un fort intérêt pour le concept d'archive ouverte surtout à cause de son accès libre et gratuit. Ces caractéristiques sont surtout reconnues comme des moyens de rendre accessible facilement et rapidement leurs publications à la communauté, et encore plus parce que TeLearn est une archive thématique. Pourtant les problèmes de droits d'auteurs qu'ils connaissent peu les freinent quelque peu dans le dépôt. D'un point de vue d'un lecteur, les chercheurs de Metah sont réticents par rapport à la qualité de ce qui est déposé. Ils ont été rassurés par le fait qu'un gestionnaire s'occupait de valider la recevabilité du contenu même sous des contraintes minimales (il n'y a pas véritablement d'évaluation de la qualité scientifique du contenu).

Les chercheurs de l'équipe Metah ont aussi proposé quelques améliorations techniques et perspectives de développement. Les améliorations techniques sont de l'ordre de la simplification des cinq étapes du processus de dépôt et de la communication entre chercheurs, par exemple en mettant au moins une adresse électronique de l'auteur ou du déposant. La connexion avec des outils de gestion bibliographique (type Endnote) leur parait indispensable. Toujours dans le souci d'apprécier la qualité des ressources, ils demandent à créer des outils autour de l'évaluation de la pertinence ; par exemple, pour constituer des comités de lecture

virtuels travaillant avec une grille de lecture en ligne ou pour noter la pertinence.

4 Perspectives de développement

4.1 Technique

Les perspectives de développement technique de TeLearn sont à réaliser sur trois niveaux :

- Les moyens d'exploitation des données ;
- Les outils de construction de communauté ;
- La pérennité des données.

4.1.1 Les moyens d'exploitation des données

Avant de mettre en œuvre ces moyens, il faudra que TeLearn gère les versions successives d'une même ressource. Cette fonctionnalité sera mise en place à la fin de l'année 2007.

Les moyens d'exploitation des ressources comprendront des **fonctionnalités d'exploration** du fonds avec un moteur de recherche en texte intégral, utilisant des opérateurs booléens, et une indexation multilingue.

De plus, des **fonctionnalités statistiques et bibliométriques** permettront d'analyser les usages et d'apprécier la pertinence de l'archive (par le facteur d'impact ou le h-index).

TeLearn proposera aussi des **fonctions d'exportation** des données sous des formats compatibles avec les logiciels bibliographiques les plus courants (par exemple, Endnote, Reference Manager, bibtext).

4.1.2 Les outils de construction de communauté

TeLearn a été conçu dans l'optique d'améliorer la communication entre les chercheurs. Des fonctionnalités de **construction de**

communauté seront développées pour donner les meilleurs moyens de réaliser cet objectif.

Ces fonctionnalités seront basées sur des outils de discussion, par exemple, des forums attachés aux publications. Mais aussi des espaces partagés permettant l'élaboration collective de référentiels (pondération par un niveau de confiance selon les angles de vue ou les démarches scientifiques ou une grille de lecture).

Une communauté pourra aussi se créer autour de la construction et du partage de bibliographies.

4.1.3 La pérennité des données

La pérennité des ressources sera assurée par les entrepôts institutionnels (notamment celui du CCSD). La synchronisation en temps réel ou un peu différé des données est prévue pour fin 2007.

4.2 De dépôt

Les perspectives de développement en terme de dépôt s'entendent pour accroître le fonds de l'archive. Plusieurs pistes peuvent être envisagées.

Tout d'abord, une **communication dynamique** et des formations seront mis en place. En effet, malgré la motivation par principe des chercheurs pour ce type d'initiative [5], le développement des pratiques et des utilisations de l'archive n'est pas aussi enthousiaste.

Une autre piste est de créer des partenariats avec les organisateurs de conférences en TEL (société savante, université, éditeur scientifique) ou de revues institutionnelles. Dans cet esprit, un partenariat a été établi en octobre 2007 pour accueillir les archives de la revue internationale Artificial Intelligence in Education ; créée en 1989, cette revue publiée par IOS Press accepte l'accès ouvert à ses publications après une période d'embargo de deux ans.

Conclusion

Au terme de cette présentation, nous avons pu voir le fonctionnement d'une archive ouverte unique. Sa spécificité tient au fait qu'elle est multilingue et centrée sur le thème de l'apprentissage par les technologies ou « Technology Enhanced Learning ». Sa pertinence grâce à son système de double modération (déposants et thématique), les services de bases proposés (actualité, liens, aide, fils RSS) et son interopérabilité commencent tout juste en moins de deux ans d'existence à agréger une communauté plus large que celle du réseau d'excellence Kaleidoscope ; comme le montre, le noyau de 437 visiteurs différents par semaine et le nombre toujours croissant de ressources. Soulignons aussi la modularité du système qui permet de créer d'autres archives ouvertes.

Malgré ces efforts, TeLearn arrive à peine au stade de la consolidation de ses services et de la constitution d'un fonds significatif qui permettront de mettre en place des services avancés de communication et de partage entre les chercheurs. Peut être que la pérennisation de l'archive avec le CCSD et une plus grande visibilité internationale à travers des projets Google Scholar et Scientific commons assureront la crédibilité de TeLearn sur le long terme.

Références

[1] ROMARY Laurent. *Préface*. In : AUBRY Christine et JANIK Joanna. Coord. Les Archives Ouvertes : enjeux et pratiques - Guide à l'usage des professionnels de l'information. Paris : ADBS, 2005, p.5-8. ISBN 2-84365-079-8

[2] FONTÈS Olivier, PUTZ Edouard. *Introduction à AWSTATS.* [en ligne]. EITIC. Disponible sur <http://www.farafina.org/alternc/aide_awstats.pdf> (consulté le 13 septembre 2007)

[3] MAGRON Agnès. *Auto-archivage des publications scientifiques Synthèse d'enquêtes menées auprès des chercheurs.* [en ligne]. Institut des Sciences de l'Homme, 2007. Disponible sur

<http://archivesic.ccsd.cnrs.fr/docs/00/15/15/75/HTML/index.html>. (consulté le 6 juillet 2007)

[4] SWAN Alma, BROWN Sheridan. *Open access self-archiving: an author study.* [en ligne]. Key Perspectives Ldt, 2005. Disponible sur <http://eprints.ecs.soton.ac.uk/10999/01/jisc2.pdf> (consulté le 6 juillet 2007)

[5] WOJCIECHOWSKA Anna. *Analyse d'usage des archives ouvertes dans le domaine des mathématiques et l'informatique*. In : Documentaliste, Sciences de l'Information, vol. 43. Paris : 2006, p.5-6.