

Retours sur l'utilisation d'une solution numérique d'aide à la détection du Plagiat : retours d'expériences, statistiques d'utilisation et comportements adaptés à son usage

Antoine Jean-Michel

Directeur de la Sous Direction des Usages du Numérique - Université de Lorraine
30 Rue Lionnois - B.P 60120 - 54003 Nancy Cedex
Tél. : 03 83 68 20 90
Jean-Michel.Antoine@univ-lorraine.fr

Fay-Varnier Christine

Maître de conférences - Chargée de mission TICE - Université de Lorraine
ENSG - rue du doyen Marcel Roubault - B.P. 40 - 54501 Vandoeuvre-lès-Nancy
Tél. : 03 83 59 64 34
Christine.Fay-Varnier@univ-lorraine.fr

Pino Pierre

Maître de conférences - Chargé de mission TICE - Université de Lorraine
Université Paul Verlaine Metz, Ile du Saulcy - BP 80794 - 57012 Metz cedex 1
Tél. : 03 87 54 70 87
Pierre.Pino@univ-lorraine.fr

Roussanaly Azim

Maître de conférences - Chargé de mission TICE - Université de Lorraine
UFR Mathématiques et Informatique - Pôle Lorrain de Gestion - 13, Rue Michel Ney
CO 75 - 54037 Nancy Cedex
Tél. : 03 54 50 35 28
Azim.Roussanaly@univ-lorraine.fr

Introduction

Le plagiat est désormais une réalité dans les travaux des étudiants et il est largement facilité par l'accès aux ressources numériques via Internet, le développement des ressources pédagogiques mises à disposition par les Universités Numériques Thématiques (UNT), les sites de ventes en ligne de rédactions, devoirs, essais, etc.

Il y a une méconnaissance assez générale des règles en la matière chez les étudiants comme chez les enseignants. Le développement de l'usage du numérique dans la pédagogie nécessite donc d'éduquer les utilisateurs aux règles d'utilisation des documents utilisés dans la réalisation des travaux. A l'université, il convient de déployer une réelle politique de sensibilisation, voire de dissuasion, par le déploiement de dispositifs d'information et de formation consolidés par des dispositifs techniques d'analyse des travaux étudiants.

Michelle Bergadaa, spécialiste de ce domaine, propose dans son "Petit manuel à l'usage des dirigeants d'établissements universitaires" [1] huit recommandations pour lutter contre la progression de ce phénomène à l'université. Elle prône en particulier dans sa première proposition que : *"Des processus de lutte intégrés contre le plagiat doivent être élaborés, couvrant les compétences informationnelles des étudiants, l'information aux enseignants et l'intégration d'outils de détections des similitudes adaptés, des procédures d'analyse des dossiers des plagieurs et d'information des sanctions relatives aux délits détectés."*

Les universités lorraines se sont ainsi équipées d'un outil de détection de similitudes et dans cette présentation nous dressons un premier bilan sur les aspects concrets de mise en œuvre de l'outil et d'accompagnement des usagers à son utilisation.

Choix d'une solution d'aide à la détection de similitudes - Démarche

Notre démarche pour l'étude et l'acquisition d'un outil de détection de similitudes a été motivée par une demande croissante de la part de nos collègues enseignants et chercheurs observant l'augmentation du phénomène du "copier-coller" et cherchant un outil les aidant à "traquer" le plagiat. Certaines composantes de l'université, UFR, département d'IUT, ou école doctorale avaient commencé à acquérir localement de tels outils.

La recherche de solution a été initiée au sein de l'UNR UNIRE qui rassemble les universités de Bourgogne, Champagne-Ardenne, Franche-Comté et Lorraine. La volonté commune des universités partenaires d'UNIRE était d'implanter une solution d'aide à la détection des similitudes associée à la mise en place d'une base commune de documents afin de disposer d'une base de données de documents de référence plus importante.

Afin d'étudier les solutions existantes sur le marché, nous avons lancé plusieurs démarches en parallèle. Nous avons d'une part recueilli le retour d'expériences des composantes déjà dotées d'un outil de détection de similitudes et d'autre part, nous avons rencontré plusieurs éditeurs de solutions qui ont présenté aux partenaires d'UNIRE leurs applications respectives.

Nous avons pu bénéficier gratuitement par la suite de comptes utilisateurs pendant une période assez longue, ce qui nous a permis de déployer des tests "grandeur nature". Un panel d'utilisateurs parmi des enseignants et des chercheurs particulièrement motivés par ce sujet ont été identifiés et ont pu contribuer à ces tests pendant un semestre complet.

Cette période d'expérimentation nous a permis d'identifier les besoins et les attentes précises et de rédiger le cahier des charges nécessaire au lancement de l'appel d'offre que nous souhaitions lancer. Dans ce cahier des charges, nous avons en particulier insisté sur l'intégration de l'outil dans le système d'information de l'établissement et notamment dans l'Environnement Numérique de Travail (ENT) et la plateforme de ressources pédagogiques en ligne.

En fonction du degré de maturité de chaque établissement par rapport à cette démarche et au vu des tarifs proposés par les différents éditeurs, seules les universités lorraines ont décidé de poursuivre la démarche et l'acquisition d'une solution d'aide à la détection de plagiat.

Les universités lorraines ont donc acquis la solution *Ephorus* et l'ont déployé dès le printemps 2010.

Implémentation de la solution

Intégration de la solution dans le SI de l'établissement

La solution de base proposée par *Ephorus* s'appuie sur des comptes utilisateurs permettant l'accès au service à partir de leur propre site web. Ce type de fonctionnement et surtout la gestion pour un ensemble important d'utilisateurs (environ 4000 comptes sur l'ensemble de l'Université de Lorraine) n'est pas maintenable facilement au sein d'un établissement. Nous avons donc fait le choix d'intégrer la solution dans le système d'information de nos établissements : ENT et plateforme de dépôt de ressources pédagogiques numériques.

Les quatre universités lorraines, désormais fusionnées en Université de Lorraine, disposaient toutes de l'ENT « ESUP Portail » et de la plateforme de dépôt de ressources pédagogiques « Moodle ». Nous avons donc développé un « canal ESUP » baptisé PACO (Paper Analyse for Control of Origin) qui permet d'accéder de manière intégrée et simplifiée à l'outil *Ephorus* : dépôt de documents pour analyse, alimentation ou non de la base établissement, consultation simple des rapports de similitudes renvoyés. Cet outil a été intégré dans les 4 ENT, ainsi, les utilisateurs peuvent accéder directement et de manière transparente, à leur espace individuel respectif au sein de la solution logicielle de détection des similitudes sans avoir à se ré-authentifier sur un site web tierce.

A Nancy, l'outil n'a été rendu accessible qu'aux enseignants, chercheurs et personnels de l'université. A Metz, l'outil était également proposé aux étudiants après que ceux-ci aient pris connaissance de la charte « Anti-Plagiat » (validée en CA en novembre 2008). Chaque étudiant peut ainsi lancer une demande d'analyse, le rapport des similitudes est alors envoyé à un enseignant dont l'adresse est précisée au moment du dépôt du document. L'enseignant peut décider de rendre visible le rapport à l'étudiant qui l'a déposé. Les universités nancéennes ont choisi dans un premier temps de ne pas laisser l'accès aux étudiants, une des raisons étant que les profils utilisés dans l'ENT ne permettaient pas de le mettre en place facilement.

Une aide en ligne a été rédigée et une procédure d'assistance en ligne a été mise en place. Celle-ci est accessible directement au sein du module PACO. On prévoit d'ajouter une FAQ prochainement.

La solution *Ephorus*, est proposée avec un plugin Moodle offrant un "dépôt de devoirs avec analyse anti-plagiat". Ce plugin a été installé et utilisé en premier lieu sur la plateforme Moodle de Metz, mais ne correspondait pas à nos attentes. Suite à nos demandes d'évolution, *Ephorus* a repris intégralement le développement de ce plugin. Nous ne disposons donc à ce jour que d'un retour d'expérience messin sur l'utilisation de cet outil intégré à la plateforme de ressources pédagogiques en ligne Moodle.

Accompagnement du déploiement de l'outil

L'information autour de la mise à disposition de cet outil d'aide à la détection de plagiat s'est effectuée par les moyens classiques mais réguliers : correspondants TICE, mailing listes personnels, newsletters des universités, etc.

Nous avons également effectué des présentations aux CEVU, CS, conseil des écoles doctorales, réunions pédagogiques, SCD, séminaires d'accueil des nouveaux arrivants, etc.

Les quelques expériences précédentes dans les composantes déjà utilisatrices de l'outil nous avaient sensibilisé sur le fait qu'un tel outil n'est **qu'un outil d'aide à la détection des similitudes** et non de plagiat, et nous avons insisté sur ce point lors de nos présentations. L'utilisation de l'outil implique **une analyse nécessaire** de la part du lecteur du rapport, il ne faut pas se contenter de regarder les taux de similitudes affichés mais effectuer **systématiquement un examen plus approfondi** et faire **appel à son sens critique** pour évaluer si il y a réellement plagiat ou non. L'outil intégré à l'ENT s'appelle d'ailleurs "**Aide** à la détection de plagiat" et non "Anti-plagiat".

La question sur le calcul du taux de similitude affiché et son interprétation est d'ailleurs une de celles qui revient de façon récurrente. Les enseignants-chercheurs se sont montrés très sensibles également quant à la confidentialité des documents analysés ou déposés pour alimenter la base établissement.

Suite à ces présentations, les SCD et les écoles doctorales élaborent progressivement des procédures afin de soumettre et enregistrer systématiquement les thèses soutenues dans l'établissement. Certains SCD intègrent la présentation de la solution lors de leur formation à la méthodologie de recherche de l'information.

En parallèle, a été mis en place un groupe de réflexion autour des aspects juridiques et des conséquences que peut avoir l'utilisation d'un tel outil.

A partir de quand peut-on décider qu'on est en présence d'un plagiat ? Quelles sanctions peuvent être prises à l'encontre d'un étudiant quand un tel constat a été fait ? Est-il indispensable de réunir la commission de discipline de l'établissement à chaque fois ? Telles ont été les principales questions soulevées lors des présentations et des réflexions associées. A l'heure actuelle, les réponses n'ont pas été toutes apportées de manière définitive au sein de l'Université de Lorraine.

Il est également important d'assurer une action de sensibilisation auprès du public étudiant. Depuis 2008, l'Université Paul Verlaine Metz a intégré une partie spécifique au plagiat dans sa charte des

examens et contrôle de connaissances ([4]). La section disciplinaire de l'ex université messine utilise le module PACO dès qu'un problème de plagiat est soulevé par un enseignant. Suite à une telle démarche, un cas de plagiat avéré a déjà été constaté et a donné suite à des sanctions.

Cette sensibilisation au plagiat a également été faite lors des sessions/semaines d'accueil des étudiants à l'université et lors des cours de formation au C2i qui sont assurés à l'ensemble des étudiants de niveau L.

Retours d'expérience

Nous présentons dans cet article notre retour d'expérience sur la première année d'utilisation de la solution de recherche de similitudes dans notre université. Dans un premier temps, nous présentons quelques failles relevées dans le fonctionnement de ce type d'outil. Puis nous commenterons les résultats d'une enquête lancée récemment auprès des usagers de l'application.

Les failles

Suite aux remontées de quelques dysfonctionnements du logiciel relevés par des utilisateurs, en particulier sur des échecs d'analyse, les ingénieurs chargés de l'intégration de l'application ont réalisé des tests systématiques sur l'outil.

Problème des homoglyphes

La première faille observée est l'attaque par homoglyphes :

- le caractère l (L) apparaît visuellement de manière identique au caractère l (i) en arial
- le caractère 1 (L) est identique au caractère l (un) en Times

Il suffit donc de remplacer certains caractères par d'autres (dans une autre police) pour que deux textes absolument identiques pour le lecteur à l'écran et à l'impression papier, paraissent en fait originaux lors de l'analyse par la solution de recherche de similitudes.

Nous avons effectué des tests en remplaçant sur un fichier 196 l (L) en l (i) en police « arial », et seulement 20% de similitude ont été détectés. En faisant des substitutions un peu plus poussées à partir de lettres grecques, le taux de similitudes trouvé est alors cette fois-ci nul.

Problème de l'encodage des documents au format pdf

Nous avons observé pour plusieurs documents soumis à analyse que certains outils de conversion au format pdf (portable document format) proposent un encodage qui ne respecte pas le standard. Ainsi, les documents générés ne sont pas correctement analysés. La solution n'arrivant pas à décoder les caractères textes composant le document, les recherches de similitudes échouent. L'encodage des documents pdf est une partie sensible qu'il va falloir prendre en compte dans la suite du déploiement de l'outil. Il va cependant être difficile de faire l'inventaire de tous les outils qui construisent des documents au format pdf et de contrôler le format des fichiers qu'ils génèrent...

Vu l'explosion combinatoire des tests à effectuer pour prouver l'originalité d'un document en prenant les différents types d'encodage en compte, une solution passant par une phase d'OCR (Reconnaissance Optique de Caractères) pourrait être une réponse à ce type d'attaque. Mais cela surcharge énormément la chaîne numérique de traitement de l'information ce qui est difficilement utilisable et non stable, la phase de reconnaissance OCR demandant une manipulation humaine non négligeable.

L'usage « savant » des TICE

Une autre faille répandue consiste à insérer des guillemets « peints » en blanc, donc invisibles à l'œil, qui signalent à la solution que le texte entre les « » est une citation. Du coup la solution peut ne pas

tenir compte de cette portion du document. Pour éviter cet artifice, il est nécessaire que la solution propose de prendre en compte les guillemets ou non en option.

D'autre part, certains documents stockent sous forme d'images une partie de leur texte, l'analyse par la solution n'a aucun effet puisque l'application n'analyse que les caractères sous forme de texte.

Traduire un article écrit en anglais puis le traduire à nouveau en français constitue également un autre contournement possible de l'outil...

Faiblesse de la base documentaire

Dans la détection des similitudes, il faut bien faire comprendre aux utilisateurs que c'est en déposant leurs documents qu'ils augmentent la base de données établissement qui sert à la détection. Les solutions commerciales s'appuient sur leurs propres bases générées en butinant le web public. Le fait d'augmenter le nombre de documents de la base établissement permet donc une détection plus fine et protège par là-même le plagié puisque ses documents vont servir de référence.

De plus, et nous l'avons expérimenté à l'Université de Lorraine, de par des conventions de partenariat entre leur composante et un/des éditeur(s), certains étudiants ont accès à des bases de données thématiques "privées", donc non accessible au public, et utilisent le « copier/coller » sans citer leurs sources... Les sources utilisées n'étant pas accessibles librement en ligne et donc pas intégrées dans la base du logiciel, la détection du plagiat passe par la mise en œuvre d'un partenariat entre l'éditeur privé et l'éditeur de la solution d'aide à la détection des similitudes. La convention doit permettre d'intégrer les documents dits « privés » dans la base de données propre à l'établissement. Les éditeurs sont plutôt d'accord sur le principe, car ils ne risquent pas grand-chose, leurs documents ne sont ni lisibles ni téléchargeables, ils ne servent qu'à alimenter la base de documents de l'établissement utilisés pour la comparaison des textes.

Retour des usagers

Afin d'obtenir un retour sur l'utilisation de l'outil, nous avons lancé une enquête en ligne auprès des utilisateurs référencés de la solution. Nous commentons ci-après quelques résultats significatifs obtenus suite à cette enquête.

Statistiques d'usage

- Nombre d'utilisateurs : 542 (soit 14,5 % des 3 722 enseignants-chercheurs et enseignants sur l'ensemble de l'université)
- Nombre total de documents analysés : 3 575
- Questionnaires remplis : 172 (taux de retour : 32%)
- Le nombre de documents déposés en moyenne est inférieur à 10 pour 70% des usagers, et seulement 5% des personnes ayant répondu à l'enquête ont déposé plus de 50 documents. On remarque donc qu'il n'y a pas encore d'utilisation massive de l'outil, le potentiel d'utilisation de l'application n'est pas encore atteint, ni en nombre d'utilisateurs ni en nombre de documents soumis. L'intégration généralisée du module de dépôt de devoirs dans la plateforme de ressources pédagogiques en ligne *Moodle* devrait permettre d'augmenter le nombre d'analyses en lançant l'analyse de documents par lots.
- Jusqu'à présent, l'outil est majoritairement utilisé dans un cadre pédagogique (environ 2/3 des réponses) plutôt que dans le cadre de la recherche.

Utilisation dans le cadre de la pédagogie

Dans un contexte pédagogique, les documents soumis à analyse sont principalement des mémoires et des rapports de stage, plutôt que des productions de type devoir ou TP.

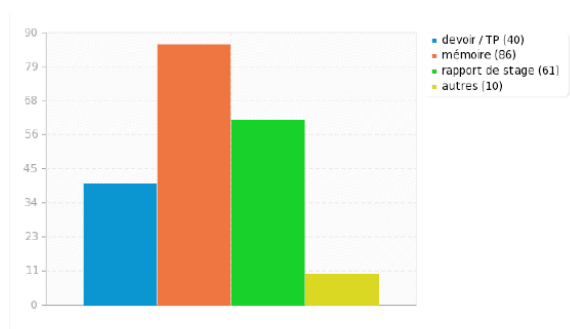


Figure 1 : Types de documents soumis dans le cadre de la pédagogie

L'outil est trop peu utilisé dans un objectif d'apprentissage à la méthodologie de recherche d'information. En revanche, pour plus de 40% des enquêtés la détection d'un plagiat potentiel a permis de déclencher des discussions avec les étudiants.

Utilisation dans le cadre de la recherche

La démarche s'avère être individuelle plutôt qu'institutionnelle et concerne principalement l'analyse de documents que l'utilisateur doit évaluer (article scientifique, articles de chercheurs encadrés, etc.). Certains enseignants-chercheurs ont utilisé l'outil pour leurs propres articles et pour détecter d'éventuels plagiats à l'encontre de leur production personnelle.

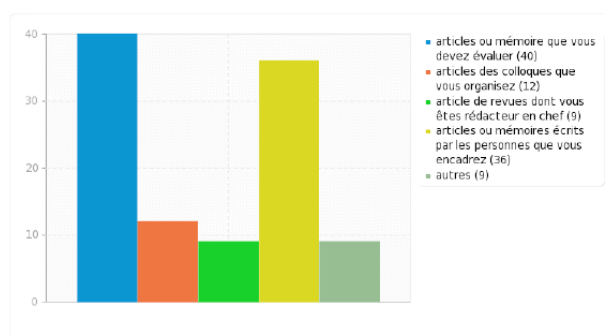


Figure 2 : Types de documents soumis dans le cadre de la recherche

- 56 % des personnes ont détecté des cas de plagiat à l'aide de l'outil.
- 80% des analyses ont été obtenues en moins de 24h, donc sauf incident technique, *Ephorus* tient ses engagements au niveau du délai de réponse.

Information et communication

Un tiers des personnes qui ont répondu à l'enquête ont eu connaissance de l'outil par le biais d'une communication ciblée : mails, présentation dans des conseils, lors de journées d'accueil, etc. C'est le bouche à oreille qui a fonctionné pour un second tiers des usagers. Les autres ont découvert l'outil plus ou moins "par hasard" en explorant les fonctionnalités de l'ENT.

On peut donc en déduire que l'information institutionnelle n'est pas celle qui est la plus efficace, et qu'il faut arriver à trouver les bons vecteurs d'information.

Interface et Formation

30% des usagers déclarent avoir éprouvé quelques difficultés dans l'utilisation de l'interface PACO accessible via l'ENT.

En analysant précisément leurs réponses, il s'avère que les difficultés rencontrées ne portent pas sur la prise en main de l'interface à proprement parler mais sur le besoin de fonctionnalités supplémentaires, sur le mode de présentation du rapport de similitude et parfois sur un échec inexplicable de l'analyse.

- Rappelons que l'application PACO mise à disposition dans l'ENT se voulait être une interface simplifiée permettant de lancer des recherches de similitudes et se chargeant de gérer la connexion vers le site d'Ephorus. Les utilisateurs un peu avertis souhaiteraient y trouver cependant des fonctionnalités supplémentaires telles que la possibilité de supprimer des documents de la base privée de l'université, de recevoir un mail prévenant de la fin d'une analyse ou celle de soumettre des documents de taille plus importante (limitation actuelle à 25 Mo).
- Au niveau de l'affichage des similitudes, certains usagers demandent à ce que celles-ci puissent être identifiées au niveau de chaque paragraphe et pas seulement à partir des sources plagiées.

Document		Rapport		
Nom du fichier	ECJS.odt	%	Report	Type
Nom	Christine fayvar36 FAY VARNIER	17%	http://fr.wikipedia.org/wiki/Pauvret%C3%A9	internet
Date	2011-04-11 22:16:00	17%	http://www.recherche.fr/encyclopedie/Pauvret%C3%A9	internet
%	29%	16%	http://www.droit-du-travail.wikibis.com/pauvrete.php	internet
		16%	http://dictionary.sensagent.com/pauvrete/fr-fr/	internet
		16%	http://dicionario.sensagent.com/Pauvrete/fr-fr/	internet
		16%	http://www.9rayti.com/uploads/content/50183cd3abffd3875d1f141c55d2a22c.doc	internet
		9%	http://desirsdavenir31100.over-blog.com/categorie-10936665.html	internet
		8%	http://www.maxisciences.com/pauvrete%9/tout-savoir.html	internet
		9%	http://histoire-geographie.ac-toulouse.fr/automne_modules_files/standard/public/p376_068ceb016b12855eca81efc59d09e9G2_-_Theme_4_-_La_pauvrete_dans_le_monde_Fabienne_GIBERT.pdf	internet
		3%	http://ca-peut-changer-le-monde.skyrock.com/	internet

▼ Allez au premier tableau de résultat

Pauvreté et marginalité sociale Introduction Beaucoup de personnes vivent dans une grande pauvreté aujourd'hui dans le monde. Cette situation les amène à vivre en marge de la société et leur exclusion fait qu'ils n'ont plus ou pas l'envie de participer activement à la société. Nous allons présenter les différentes causes de pauvreté dans le monde, ce que cela implique pour les personnes dans cette situation et en particulier la marginalité sociale qui les touche. I. La Pauvreté

▼ La **pauvreté** est l'insuffisance de ressources matérielles, comme la nourriture, l'accès à l'eau potable, les vêtements, le logement, et des conditions de vie en général, mais également de ressources intangibles comme l'accès à l'éducation, l'exercice d'une activité valorisante, le respect reçu des autres citoyens. Elle touche des personnes isolées ou des groupes et populations entières dans les pays développés, voire la majorité de la population dans certains pays en développement. Des analyses économiques et des débats portent sur la mesure de la pauvreté, ses

La **pauvreté** est l'insuffisance de ressources matérielles, comme la nourriture, l'accès à l'eau potable, les vêtements, le logement, et des conditions de vie en général, mais également de ressources intangibles comme l'accès à l'éducation, l'exercice d'une activité valorisante, le respect reçu des autres citoyens. Elle touche des personnes isolées ou des groupes et populations entières dans les pays développés, voire la majorité de la population dans certains pays en développement. Des analyses économiques et des débats portent sur la mesure de la pauvreté, ses

Figure 3 : Aperçu de le l'affichage des taux de similitudes détectés

- La dernière réserve exprimée porte sur les résultats obtenus. On constate que les usagers ont éprouvé le système et ont plus ou moins essayé de le mettre en défaut. Ils ont pu alors faire remonter certains cas de plagiat non détectés par la solution *Ephorus*, alors qu'une recherche sur google le permet. L'échec de ces analyses amène les usagers à s'interroger sur la construction de la base servant à la comparaison.

Seulement 40% des enquêtés apprécieraient avoir une information un peu plus complète sur les taux de similitudes affichés, alors que c'était au départ une interrogation forte lors des présentations qui ont été faite de l'outil.

Taux de satisfaction

80% des personnes ayant répondu à l'enquête **recommanderaient l'outil** à leurs collègues. On peut donc conclure que globalement, le service apporté par l'outil est satisfaisant et répond à un besoin des enseignants-chercheurs.

Utilisations détournées

Les usagers ayant répondu à l'enquête ne font pas état d'une utilisation détournée de l'outil eu égard à son objectif initial.

En revanche, nos services de communications institutionnels commencent à s'en servir pour évaluer l'impact de leurs communiqués sur le web. En effet suite à la diffusion d'une lettre d'information à l'extérieur de l'université, il est très difficile d'évaluer la portée de ses écrits. L'utilisation à contre-courant de la solution est pour le coup utile et rapide. Le pourcentage de similitudes indique l'impact de diffusion, et pour ce type d'utilisation, plus le résultat est élevé, mieux c'est !

Conclusion et perspectives

Evolution des outils

Par la mise en place d'une solution de détection des similitudes, notre objectif est essentiellement pédagogique : apprendre aux étudiants à ne pas plagier, à respecter le droit à la propriété intellectuelle (citer ses sources,...).

Quelque soit la solution installée, on rencontrera des cas où la recherche échouera et ne trouvera pas de documents similaires, ou bien un nombre réduit de documents similaires. En effet, les documents qui ne sont pas dans la base ne peuvent pas être utilisés pour l'analyse, et sauf à être aussi puissant que *google*, les bases utilisées seront toujours difficilement à jour...

De plus, il suffit qu'un étudiant copie une partie d'un livre qui n'est pas numérisé, qu'il utilise un texte extrait d'un site web non référencé (base privée) ou mette en œuvre un des contournements décrits précédemment (ou d'autres [3]) pour que les similitudes ne soient pas diagnostiquées.

Il est certain que les applications du marché vont être améliorées au fur et à mesure pour palier ces faiblesses, mais elles risquent d'intervenir souvent après coup.

Une première solution consisterait en premier lieu à n'autoriser que le dépôt des fichiers texte "brut", cela permettrait d'enrayer grandement ces «contournements». Aujourd'hui, avec un peu de connaissances TICE bureautique (C2i), tout un chacun est capable d'exporter un document mis en forme dans un traitement de texte en un texte brut, sinon l'application peut aider à le faire. L'outil d'analyse utiliserait alors uniquement le texte du document, le traitement en serait allégé et les biais observés lors de l'utilisation de l'outil seraient réduits.

Le paradoxe : publier pour ne pas être plagié

Par ailleurs, nous avons affaire à un paradoxe : « publier massivement sur le web les travaux produits à l'université (étudiants et enseignants-chercheurs) pour mieux détecter des similitudes ».

Il faut que nous incitions les étudiants et enseignants-chercheurs à déposer leurs travaux dans les bases des établissements pour pouvoir les prendre en compte lors de la recherche de similitude.

Certaines personnes ont peur de mettre leurs travaux sur internet de peur d'être plagiées, alors qu'au contraire avec le développement de ces outils, il serait préférable de mettre les travaux sur internet pour que les similitudes soient détectées.

Former et informer avant de sanctionner

En conclusion, d'après notre expérience, nous sommes très favorables à l'utilisation d'une solution logicielle pour gérer la masse importante de documents venant des filières nombreuses et variées.

Mais il est indispensable de mettre en place des actions de prévention/formation avant d'utiliser cette solution et avant d'appliquer toute forme de sanction. Un plan de formation et de sensibilisation à l'honnêteté intellectuelle doit être systématiquement mis en place en parallèle au déploiement d'une solution de recherche de similitudes.

Le simple fait d'annoncer que l'établissement dispose d'une solution de détection des similitudes en vue d'aider les enseignants à détecter les plagiats est déjà une première étape, en quelque sorte c'est la « peur du gendarme »...

Nous devons pour ce faire, accentuer encore plus notre collaboration avec les SCD.

D'autre part, pour gérer finement un document, l'utilisation seule de la solution ne suffit pas. Les résultats rendus par l'application ne fournissent qu'une indication. Une action de sensibilisation auprès des enseignants est également indispensable, il faut insister sur le fait qu'un travail d'analyse important doit être assurée par les enseignants ou chercheurs afin dans de décider s'il y a vraiment plagiat avéré ou pas et ne pas prendre de façon brute le taux de pourcentage de similitudes détecté par l'application.

Aspects juridiques et sanctions

Nous attendons la mise en place des conseils de la toute nouvelle Université de Lorraine pour poursuivre nos travaux autour des aspects juridiques relevant des problèmes de plagiat. Les usagers sont en attente d'aide sur ces aspects et sur les différents critères qui amènent à suspecter un plagiat. Nous envisagerons également d'accompagner les actions de sensibilisation et de formation par la mise en place d'une charte de "respect de l'éthique et de la déontologie d'internet".

Extensions de la base privée

Si les autres universités d'UNIRE envisagent d'acquérir cette solution, nous envisagerons des accords permettant d'élargir la base privée à une base privée UNIRE.

Une autre perspective intéressante serait d'élargir la base privée en y ajoutant des accès aux ouvrages et revues auxquelles l'université est abonnée. Pour ce faire, il est nécessaire de négocier des accords avec les éditeurs. Des démarches au niveau national seront certainement plus efficaces que des celles qui pourraient être menées au niveau régional.

Même si la solution mise en œuvre et si les outils disponibles ne sont pas encore parfaits et exempts de tout reproche et de « faux négatifs », notre expérience s'avère très positive et le retour des utilisateurs satisfaits est très encourageant, il nous incite à poursuivre notre démarche et nos efforts. La généralisation et la globalisation de cette démarche ne pourront qu'accentuer les résultats obtenus de manière à ce que l'ensemble des acteurs du monde universitaire intègrent cette problématique dans leur quotidien.

Référence

- [1] http://responsable.unige.ch/Documents/Manuel_usage_dirigeants.pdf
- [2] http://responsable.unige.ch/Documents/masterplagiat_V3.pdf
- [3] <http://archeologie-du-copier-coller.blogspot.fr/2010/01/xxxqxqxqsucasuca.html>
- [4] http://www.univ-metz.fr/vie_etudiante/droits-devoirs/Charte-approuvee-par-CA-du-18-11-02.pdf