

## FDP 2007 « plagiat » - Annexe 1

### Rapport de l'expérimentation de Compilatio

#### **Introduction**

Suite aux travaux menés durant l'année 2006-2007, le groupe de travail lié au FDP « plagiat », appuyé par le CEFO, a mis sur pied une expérimentation d'un logiciel de détection du plagiat, durant les sessions de juin et de septembre 2008.

Le groupe souhaitait expérimenter

- Un **système institutionnel de détection**, c'est-à-dire installé sur un serveur, de manière à être accessible à l'ensemble des enseignants, par opposition à un logiciel que chaque enseignant doit télécharger sur sa propre machine. Un tel système institutionnel présente les avantages d'une part, d'offrir un service centralisé à l'ensemble des enseignants (optimisation de la gestion des ressources pour le plus large public possible), et d'autre part, de générer une base de données sur l'ensemble des travaux rédigés dans l'institution (plus la base de données est grande, plus les plagiat entre travaux sont détectables).
- Un **système de détection intégré à iCampus**. En effet, de nombreux enseignants utilisent iCampus pour réceptionner la version électronique des travaux des étudiants. Leur permettre, par un simple clic, d'analyser ces travaux et de voir les résultats directement dans iCampus représente la solution la plus simple et la plus efficace, le principe d'économie représentant une condition nécessaire pour l'intégration efficace d'un nouvel outil dans les pratiques quotidiennes (Rabardel, 1995<sup>1</sup>).

#### **Choix du logiciel de détection**

Suite à une analyse de l'offre en matière de solutions de détection du plagiat, des contacts ont été pris avec deux sociétés : Compilatio et Ephorus. Le choix s'est porté sur Compilatio : celui-ci présentait l'avantage unique de posséder un plug-in qui permettait facilement (selon la société Compilatio) de l'intégrer dans la plate-forme Claroline. Pour tout autre logiciel, il aurait fallu consacrer du travail d'informaticien pour programmer un tel plug-in.

Une licence d'utilisation de Compilatio a été achetée, couvrant une année d'utilisation (mai 2008 à avril 2009) et prévoyant 1000 crédits d'analyse (2300 euros, *in fine*). L'installation du plug-in a été réalisée par l'équipe iCampus.

En mettant en avant l'existence du plug-in d'intégration à Claroline au moment de l'appel d'offre, la société Compilatio a fait jouer un argument marketing qui s'est révélé décevant. Un plug-in avait effectivement été développé, mais jamais testé. L'intégration à iCampus a représenté le premier test « grandeur nature » et a révélé de nombreux problèmes dans le code informatique du plug-in. L'équipe iCampus a dû beaucoup retravailler ce code, en collaboration avec les informaticiens de Compilatio. Aucune des deux équipes n'ayant planifié ce travail inattendu, l'intégration à iCampus a été longue et laborieuse, et des bugs ont persisté jusqu'à la fin de l'expérimentation, pour certains enseignants. La société Compilatio a réduit sa facture de moitié pour cette raison.

---

<sup>1</sup> Rabardel, P. (1995). *Les hommes et les technologies, approche cognitive des instruments contemporains*. Paris : Armand Colin.

## Description de l'échantillon

L'échantillon a été constitué sur une base volontaire. Un appel à participation a été envoyé par email, en février 2008, à 49 membres de l'UCL qui avaient participé aux séances de travail lors de l'année académique 2006-2007 ou avaient manifesté leur intérêt autour de cette problématique. Ils étaient invités à relayer l'invitation autour d'eux. Par ailleurs, l'appel à participation a été diffusé via Résonances (avril 2008), dans La Quinzaine (avril 2008), via le site web de l'IPM et via les pages « plagiat » sur le portail de l'UCL.

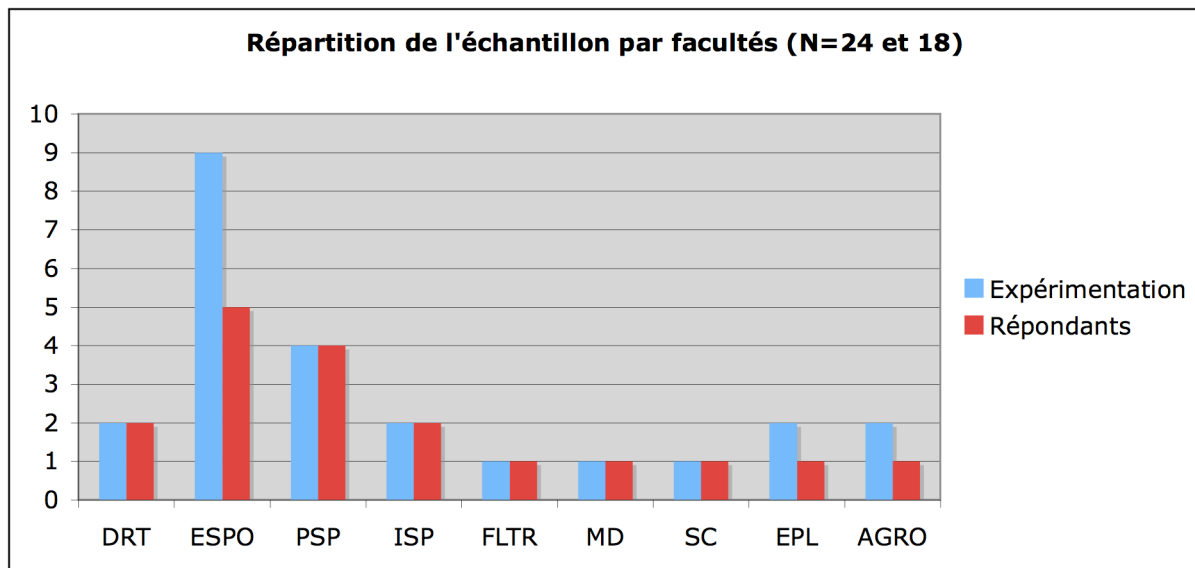
24 enseignants ont répondu positivement à l'appel. Ils ont bénéficié du « plug-in » Compilatio dans un ou plusieurs espaces iCampus qu'ils nous ont indiqués. Ce plug-in a été effectif de début mai à mi-septembre (jusqu'à la mise à jour d'iCampus pour la nouvelle année académique), pour deux sessions d'examens, donc.

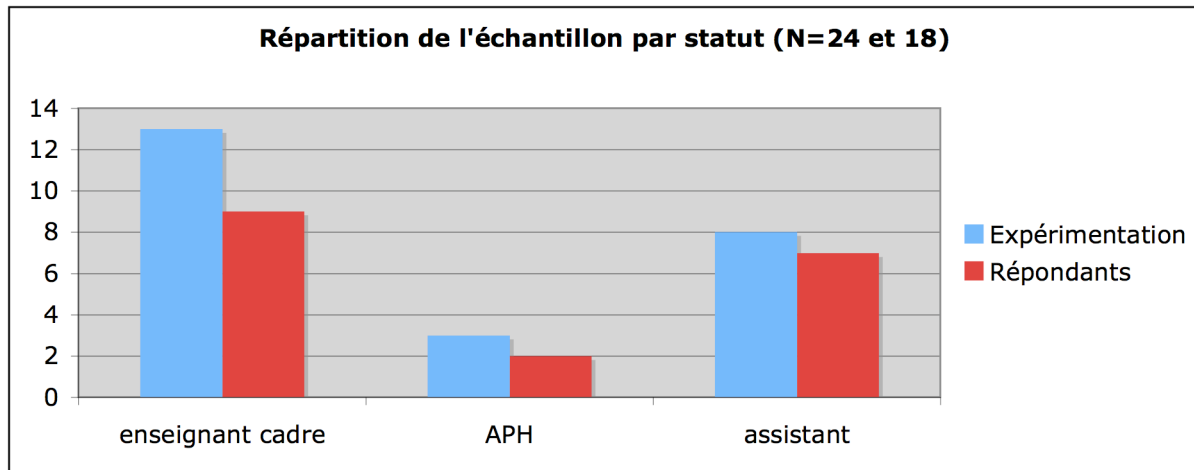
Sur ces 24 enseignants,

- 5 n'ont vraisemblablement pas utilisé Compilatio : ils ont soit répondu négativement à l'enquête intermédiaire, soit pas répondu à aucune des deux enquêtes.
- 18 ont répondu à l'enquête finale, organisée en ligne durant le mois d'octobre 2008.

Sur les 18 enseignants qui ont répondu à l'enquête finale, 3 n'ont pas réussi à analyser des documents par Compilatio : 2 ont expérimenté des problèmes techniques qui n'ont pas pu être résolus ; pour le troisième, le type de travaux fournis par les étudiants n'était pas compatible avec Compilatio (plusieurs fichiers générés par LaTeX et zippés ensemble).

L'échantillon initial (expérimentation) et final (répondants à l'enquête) se répartit comme suit :





Il faut remarquer que, par rapport au nombre d'enseignants qui avaient manifesté leur intérêt et exprimé l'importance qu'ils accordaient à la problématique du plagiat, l'échantillon de volontaires est finalement assez réduit. Nous espérions constituer un échantillon d'une cinquantaine de participants, et nous n'en sommes arrivés qu'à 24, malgré la grande couverture de communication. Ce fait nous semble important à prendre en compte dans la réflexion sur les suites à donner à cette expérimentation : y a-t-il réellement un grand nombre d'enseignants désireux d'utiliser un logiciel de détection ?

### ***Description de l'enquête de fin d'expérimentation***

Une invitation à répondre à un questionnaire en ligne (LimeSurvey) a été envoyée aux 24 enseignants de l'échantillon le 10 octobre. L'enquête est restée ouverte jusqu'au 24 octobre, un rappel par email ayant été envoyé le 21.

Ce questionnaire comprenait 31 questions, réparties en sections comme suit :

- Bilan de l'utilisation de Compilatio et des cas de plagiat détectés
- Appréciation de Compilatio
- Intérêt pour l'UCL de se doter d'un logiciel de détection (questions ouvertes)
- Appréciation de différents scénarios possibles d'utilisation d'un logiciel de détection à l'UCL (questions fermées, réponses à l'aide d'une échelle de Likert à 4 degrés)
- Suggestions d'actions de prévention du plagiat à l'UCL

### ***Analyse des résultats***

#### **1. Utilisation de Compilatio durant les deux sessions d'examens**

350 documents, dont 12 mémoires, ont été analysés par les 15 enseignants qui ont finalement utilisé Compilatio. Sur les 1000 crédits d'analyse achetés pour la licence, 330 ont été utilisés<sup>2</sup>. En moyenne, un enseignant a analysé 23 documents pour l'ensemble des deux sessions d'examens.

4 enseignants de l'échantillon déclarent avoir détecté des cas de plagiat problématiques, dont le total se monte à 10 pour les 350 documents analysés. En réaction à ces cas de plagiat, les enseignants citent :

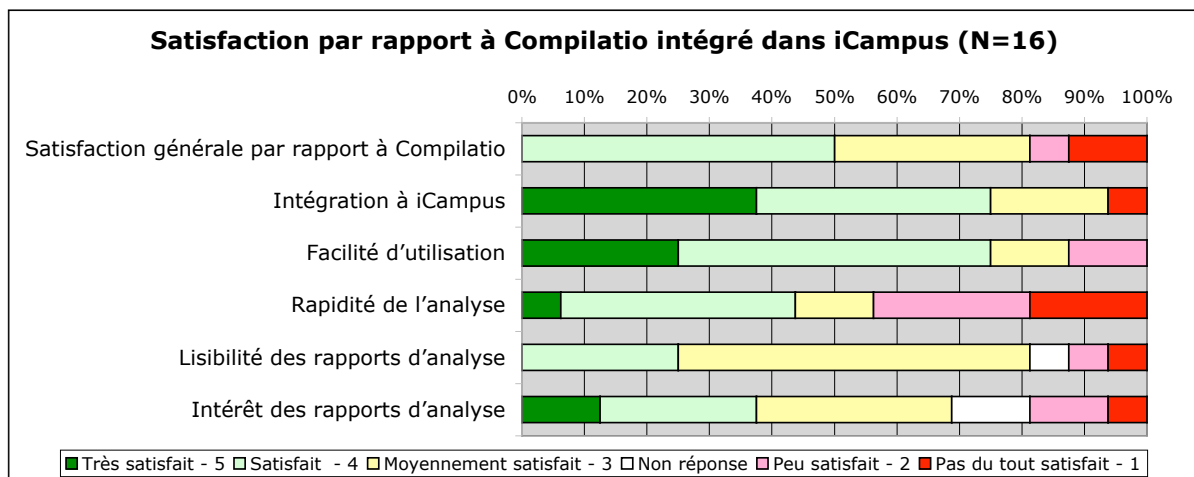
---

<sup>2</sup> Il faut savoir que le système de calcul de crédits d'analyse dans le logiciel Compilatio n'a pas fonctionné. C'est a posteriori, après la fin de l'expérimentation, que la société Compilatio s'en est rendu compte et a effectué une estimation des crédits utilisés.

- « J'ai convoqué les étudiants concernés en leur laissant la possibilité d'avouer d'eux-mêmes le plagiat. Ce qui fut le cas pour les deux étudiants en cause. J'ai veillé à m'enquérir des raisons de ce plagiat. Les deux cas étaient justifiés par un manque de temps (et non par une volonté manifeste de frauder). (...) De commun accord avec cet étudiant et compte tenu de sa volonté d'amendement, je lui ai mis 0/20 pour son travail. (...) Compte tenu de ce qu'il se trouvait en première année et de sa volonté d'amendement, j'ai préféré faire preuve de mansuétude et lui ai accordé un délai d'une semaine pour corriger son travail et me remettre un travail exempt de toute fraude (ce qui constituait en soi une "sanction", ce délai prenant place en pleine session d'examens de juin). L'étudiant a respecté ces modalités et a finalement remis un travail exempt de tout plagiat. »
- « La cote a été très fortement baissée, mettant les étudiants concernés en échec. Lors de l'explication des cotes obtenues, il leur était expliqué que le plagiat ne peut être accepté à l'université. Ces étudiants ont recommencé leur travail pour la session de septembre en tenant compte de ça. Ils ont finalement réussi. »
- « Discussion étudiant + note 0 sur le travail concerné ».
- « Entretien avec l'étudiant sur cette question qui résultait en l'obligation de recommencer le travail réalisé. Les cas de plagiat étaient essentiellement liés à des cas de méconnaissance des pratiques (étudiants étrangers) ou une absence de formation à la (bonne) recherche documentaire. »

## 2. Appréciation de Compilatio

Les 6 items suivants étaient proposés, par rapport auxquels les participants étaient invités à indiquer un niveau de satisfaction à l'aide d'une échelle de Likert à 5 degrés (pas du tout satisfait – tout à fait satisfait).



La satisfaction générale est moyenne ou bonne pour 80% des répondants (13 sur 16). L'intégration à iCampus et la facilité d'utilisation sont les items qui remportent la meilleure satisfaction (satisfaction positive pour 12 participants sur 16). Les insatisfactions liées à la rapidité d'analyse doivent être mises en lien avec les nombreux problèmes techniques qu'il y a eu en début d'expérimentation. Quant à la qualité des rapports d'analyse, elle est inhérente au choix de Compilatio.

Ce graphique confirme l'analyse a priori de l'équipe projet : Compilatio n'est pas un logiciel idéal mais il présente l'avantage précieux de pouvoir être intégré à iCampus.

Les commentaires ouverts font référence à des problèmes techniques lors de la soumission de fichiers (formats acceptés, lenteur d'upload vers le serveur...) mais surtout à des faiblesses au niveau de l'ergonomie de l'outil : vocabulaire confus au niveau des boutons d'action, boutons en anglais, manque de feed-back informant sur l'avancement du processus d'analyse, manque de feed-back en cas d'erreur du processus (on ne sait pas pourquoi cela n'a pas fonctionné)... De tels problèmes concernent l'articulation entre Compilatio et Claroline et pourraient être

améliorés avec une réécriture informatique du plug-in, dans une collaboration étroite entre les informaticiens de Compilatio et les informaticiens de Claroline.

D'autres commentaires font référence au fonctionnement de Compilatio lui-même : taille (trop petite) des fichiers acceptés, complexité de la lecture des rapports d'analyse... Nous dépendons là des améliorations que la société Compilatio voudra bien mettre en œuvre au fil de l'évolution de son logiciel.

Enfin, certains commentaires font référence au processus même de détection du plagiat : ce n'est pas le plagiat qui est détecté mais les similitudes, les sources protégées (articles de périodiques accessibles sur abonnement) ou les sources papiers (non disponibles sur internet) ne sont pas repérées, de même que les traductions de l'anglais... Ces limites ne seront jamais contournées, elles sont inhérentes à tout logiciel de détection.

11 commentaires relatifs à l'intégration dans iCampus sur 14 sont en faveur de cette intégration. Parmi les plus explicites, citons :

- « C'était très pratique ! je crois que si vous dites à un prof (ou assistant) "va te connectez à tel endroit" en plus, il risque de ne pas le faire... Il ne faut pas oublier que la qualité d'un tel programme est aussi inversement proportionnelle au temps d'adaptation nécessaire. »
- « C'est nécessaire. Si l'on doit passer d'un logiciel à un autre, cela devient fastidieux. »
- « Cela incite à utiliser Compilatio; les données sont plus facilement accessibles. C'est un réel "plus". »
- « Les avantages sont indéniables en termes de facilité d'utilisation et de gain de temps (comment trouver le temps d'uploader, un à un, 60 travaux sur le logiciel externe?). De toute évidence, le meilleur système compte tenu du nombre de travaux à corriger. »

Un seul participant souhaite un logiciel indépendant d'iCampus. Il faut savoir que cet enseignant utilisé déjà un logiciel « personnel » (installé sur son propre ordinateur), et en est satisfait.

Deux commentaires suggèrent l'intérêt d'avoir un système disponible à la fois sur iCampus et à la fois en dehors : pouvoir uploader quelques documents que les étudiants n'ont pas posté sur iCampus (mémoires, travaux isolés...).

A la question « Si l'UCL est appelée à utiliser un logiciel de détection du plagiat après cette expérimentation, lui recommandez-vous de continuer avec Compilatio intégré à iCampus ? », 8 participants répondent « oui » (soit 50%), 3 répondent « non », et 7 ne donnent pas de réponse. Les commentaires ouverts souhaitent, dans le cas d'une poursuite avec Compilatio, qu'il y ait des améliorations fonctionnelles et techniques. Un seul commentaire suggère un autre logiciel (Eve2), mais qui ne permet pas une solution institutionnelle de détection (logiciel personnel).

### **3. La détection du plagiat à l'UCL**

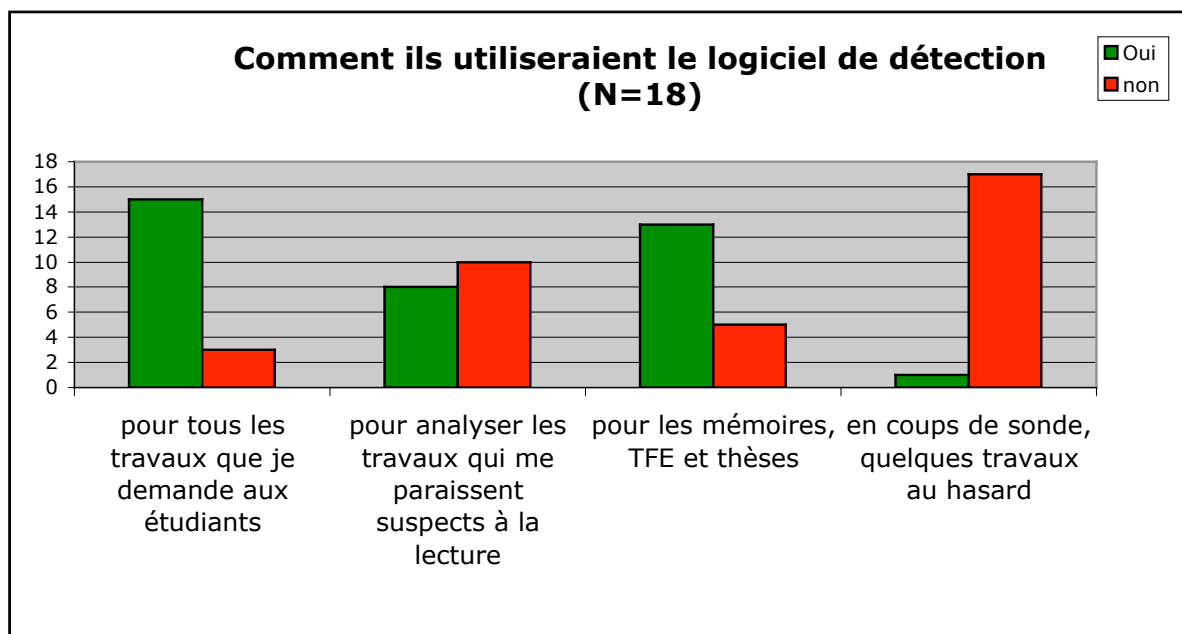
#### **A) Questions ouvertes**

A la question ouverte « Pensez-vous que l'UCL doit se doter d'un logiciel de détection du plagiat ? », 16 participants répondent par un « oui » franc. Les deux autres expriment des doutes.

- « Cela renforce l'efficacité du contrôle opéré par les enseignants et envoie un message aux étudiants. »
- « Oui, même si cela ne suffit pas. »
- « Le seul fait d'annoncer un contrôle réduit les infractions. »
- « Oui, car l'UCL ne sensibilise pas suffisamment les étudiants à ce problème. Si les enseignants utilisent ce type de logiciel, ils doivent selon moi en informer leurs étudiants, et au moins, cela aura le mérite de mettre le problème à l'ordre du jour et cela permettra à chacun d'être concerné par le plagiat. »

- « C'est indispensable. Il faut évidemment former les étudiants mais il faut aussi avoir un outil de contrôle pour montrer que vraiment on attache une grande importance à l'honnêteté intellectuelle des étudiants. »
- « Je pense qu'il s'agit d'un outil dissuasif et à la fois répressif. Si l'on rend systématique l'envoi d'une version électronique pour les travaux et mémoires, l'utilisation de Compilatio est peu coûteuse en temps, mais elle permet de détecter au minimum des cas de plagiat très clairs et d'en démontrer facilement le caractère frauduleux. Cela amènera les étudiants à veiller davantage à citer leur source et à ne pas plagier. »
- « Je ne suis pas complètement convaincu. Il y a beaucoup de moyens de rendre la détection inefficace. On risque d'entrer dans une course aux armements (détection / parade). De plus, pour les profs, c'est un truc de plus à gérer. »

Si l'UCL disposait en permanence d'un logiciel de détection, comment l'utiliseraient-ils ? Quatre propositions étaient suggérées, avec possibilité de répondre par oui ou non et de laisser des commentaires.



Commentaires :

- « Ce qui est rapide et pratique peut être largement utilisé. »
- Analyser les mémoires et les thèses « par cohérence avec le niveau d'exigence », « indispensable, étant donné l'enjeu ».

## B). Appréciation de différents scénarios de détection

Différents scénarios d'utilisation d'un logiciel de détection dans une université ont ensuite été proposés, par rapport auxquels les enseignants étaient invités à donner leur avis à l'aide d'une échelle à 4 niveaux :

- 1 – scénario exécrable pour l'UCL – à rejeter
- 2 – scénario peu intéressant pour l'UCL
- 3 – scénario intéressant pour l'UCL
- 4 – scénario excellent pour l'UCL – à retenir

Ces scénarios sont soit issus de démarches mises en œuvre dans d'autres universités, soit ont été entendus lors de rencontres ou de discussions, ou lus sur le web.

Scénarios relatifs au lieu d'installation du logiciel de détection :

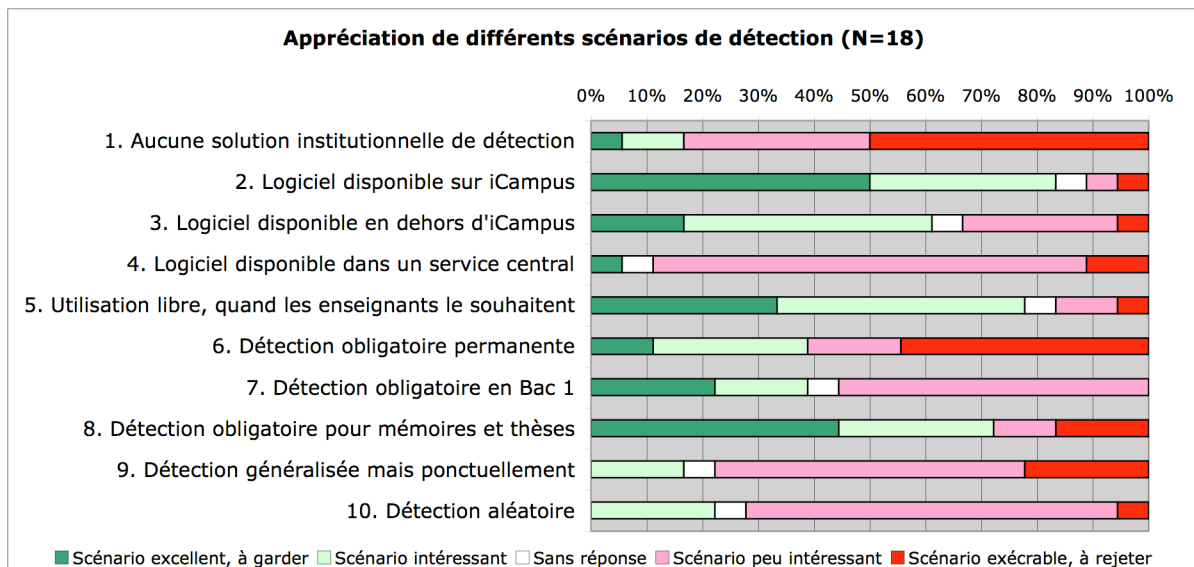
- Scénario 1 : Aucune solution institutionnelle de détection n'est prévue : les enseignants peuvent, à titre individuel (ou par département, faculté...) acheter une licence pour l'un ou l'autre logiciel de détection de leur choix.

- Scénario 2 : Un logiciel de détection est à disposition sur iCampus (comme dans cette expérimentation)
- Scénario 3 : Un logiciel de détection est disponible en dehors d'iCampus pour les enseignants qui le souhaitent (l'UCL a payé une licence pour un service de détection en ligne, par exemple)
- Scénario 4 : Les enseignants envoient les travaux problématiques à un service central de l'UCL, qui réalise l'analyse pour eux (par exemple, l'IPM, les bibliothèques...)

Scénarios relatifs à la manière de l'utiliser :

- Scénario 5 : Utilisation libre, par les enseignants qui le souhaitent quand ils le souhaitent.
- Scénario 6 : Détection obligatoire permanente : tous les enseignants doivent toujours analyser tous les travaux des étudiants.
- Scénario 7 : Détection obligatoire en Bac 1 : obligation pour les enseignants de Bac 1 d'analyser tous les travaux (pour faire prendre de bonnes habitudes aux étudiants dès le début des études)
- Scénario 8 : Détection obligatoire pour les mémoires et les thèses : obligation pour les promoteurs d'analyser tous les mémoires et toutes les thèses de doctorat (pour ne pas laisser sortir des étudiants sur base de travaux plagiés)
- Scénario 9 : Détection systématique mais ponctuelle : par exemple tous les 2-3 ans, pour une session d'examen donnée (janvier ou juin), il est demandé aux enseignants d'analyser tous les travaux réalisés lors de cette session. Ceci est assorti d'un effet d'annonce aux étudiants « pour cette session, tous vos travaux seront analysés par un logiciel de détection ».
- Scénario 10 : Détection aléatoire : un service de l'UCL (l'IPM, les bibliothèques...) soumet l'un ou l'autre travaux d'étudiants, choisis au hasard, lors de chaque session d'examen. Ceci est largement communiqué aux étudiants, pour un effet dissuasif de type « radar ».

Les appréciations se distribuent comme suit :



Les scénarios les mieux appréciés sont ceux de l'expérimentation : un logiciel de détection disponible sur iCampus, pouvant être utilisé selon les souhaits des enseignants. On peut remarquer un intérêt manifeste pour une détection obligatoire dans les travaux de mémoires et

de thèses. Les scénarios contraignants (détection obligatoire) sont peu appréciés, de même que les scénarios « organisés » au niveau institutionnel : détection systématique dans toute l'université lors de certaines sessions d'examen, ou détection aléatoire par un service central.

Commentaires des répondants :

A propos du scénario n°4 :

- « L'hypothèse n° 4 risque de se révéler extrêmement chronophage pour les personnes qui seraient chargées de cette mission, compte tenu du nombre de travaux réalisés chaque année à l'UCL. En outre, toujours dans cette hypothèse, le fait pour l'enseignant de devoir poser des démarches "administratives" (fussent-elles informatisées) risque de constituer un frein à l'utilisation. »
- « Un service pour les travaux problématiques (cfr hypothèse 4) doit être un minimum. »

A propos du scénario n°7 :

- « Ce scénario "efface" la complexité des trajectoires universitaires. Que fait-on des étudiants qui ne passent pas en BAC1 à l'UCL : FOPES, Institut des sciences du travail, passerelles des hautes écoles, et demain, master issus d'autres universités. »

A propos du scénario n°10 :

- « L'hypothèse 10 est très intéressante mais doit se faire en quantité suffisante, car si on ne scanne que 0.5% des travaux, autant dire que cela ne sert à rien. »
- « Je ne crois pas à l'effet dissuasif du logiciel (et surtout pas s'il n'est utilisé que pour des coups de sonde dont les étudiants sont avertis), mais bien à l'effet dissuasif du fait que le plagiat (détecté ou non par Compilatio) soit sanctionné à l'UCL. »

D'autre part, plusieurs commentaires insistent sur l'importance du caractère collectif de la mobilisation contre le plagiat :

- « Il me paraît délicat d'imposer une détection systématique, car je pense qu'elle doit être précédée d'avertissements clairs aux étudiants. Or, tous les enseignants ne sont pas nécessairement sensibilisés à la problématique. Il faudrait d'ailleurs peut-être commencer par cela ! »
- « Chaque enseignant devrait scanner un taux de 25% des travaux qu'il reçoit, au grand minimum. Cela ne demande que très peu de temps et permettrait d'impliquer l'ensemble du corps professoral dans l'ensemble des facultés. Si ce sont toujours les quelques mêmes qui s'épuisent à tenter de sensibiliser les étudiants, cela ne servira pas à grand chose. »
- « Il faut que les enseignants prennent en main ce problème et comprennent bien que 1/ le plagiat est très répandu 2/ il est indispensable que les étudiants comprennent que le plagiat est inadmissible. »

## **Suggestions d'action de prévention**

Trois questions ouvertes visaient à recueillir les idées des répondants à propos d'actions de prévention à destination des étudiants, des enseignants et de l'UCL dans son ensemble.

La plupart des réponses encouragent les lignes d'action déjà mises en oeuvre : campagne de sensibilisation par voies d'affiche, pages web d'information, envoi de courriers circulaires, information orale dans les auditoriums, utilisation d'un logiciel de détection... Certaines réponses suggèrent des actions de sensibilisation spécifiques en facultés.

Plusieurs commentaires signalent l'importance de sensibiliser et d'informer les enseignants, et d'attirer leur attention sur le plagiat qu'ils commettent peut-être eux-mêmes.

Sensibiliser les étudiants au fait que « l'université n'est pas contre la simple "compilation" de sources, pourvu que ces sources soient citées et que le texte soit reformulé. Un étudiant peut s'inquiéter d'avoir du mal à inventer quelque chose de nouveau, et ça peut l'amener à s'approprié (à voler) les idées d'autrui. »

Parmi les réponses, la question des sanctions est celle qui revient le plus souvent (10 allusions), d'une part dans le souhait que ces sanctions soient plus clairement annoncées aux étudiants et suggérées aux enseignants (« donner aux enseignants de ligne directrices de sanctions », « l'UCL devrait prévoir des balises claires en matière de sanctions », « harmoniser les sanctions », « ce n'est pas le système de détection mais la sévérité de la sanction qui explique le mieux l'"honnêteté" des étudiants »), d'autre part dans le souhait que les sanctions soient réellement appliquées en cas de plagiat avéré.

Enfin, un commentaire souligne que le problème est déjà présent dans les études secondaires, et qu'il serait utile de mettre un logiciel de détection gratuitement à disposition des écoles secondaires.

## **Conclusions**

18 enseignants concernés par la problématique du plagiat et volontaires pour expérimenter Compilatio intégré à iCampus ont répondu à cette enquête. Ceci représente un échantillon fort petit et non représentatif des enseignants de l'UCL, car biaisé dans sa composition. Il faut remarquer que cette invitation à utiliser un logiciel de détection pendant deux sessions d'examen n'a pas recueilli un plus grand nombre de volontaires.

Ces enseignants estiment qu'il est indispensable que l'UCL soit dotée d'un logiciel de détection du plagiat, intégré à iCampus (mais accessible également en dehors d'iCampus), que les enseignants pourraient utiliser quand et comme ils le souhaitent. Ils sont favorables à ce que des lignes directrices soient données pour détecter le plagiat de manière systématique dans les mémoires et les thèses. Les autres suggestions d'actions de détection systématiques ou contraignantes sont moins appréciées.

Compilatio est globalement apprécié, même si des problèmes techniques sont à résoudre et si l'ergonomie du logiciel est à améliorer. Si l'UCL envisage de continuer avec ce logiciel, il faut prévoir du temps de travail d'informaticiens pour élaborer, en concertation avec la société Compilatio, une nouvelle version du plug-in parfaitement adaptée à Claroline. L'UCL doit également prévoir un budget annuel d'achat de la licence de Compilatio.

Si l'on doit envisager de (re)programmer un plug-in d'intégration à iCampus, il vaut la peine de reprendre contact avec d'autres logiciels institutionnels de détection, comme Ephorus, et de voir s'ils ne présentent pas des rapports d'analyse plus lisibles que ceux de Compilatio.

Une alternative à l'achat d'un logiciel commercial est de développer un logiciel maison, selon une licence Open Source, qui puisse être distribuée gratuitement aux écoles secondaires. Cette alternative est discutée en annexe 2.

Parmi les actions de prévention suggérées par les répondants, des lignes directrices, balises claires, en matière de sanction sont souhaitées, ainsi qu'une sensibilisation des enseignants à l'importance de sanctionner effectivement les cas de plagiat.