



Apports de l'application web libre MarkUs dans le processus de correction de travaux en informatique

► Morgan MAGNIN (Ecole Centrale de Nantes – IRCCyN, Nantes),
Guillaume MOREAU (Ecole Centrale de Nantes – CERMA, Nantes),
Vincent TOURRE (Ecole Centrale de Nantes – CERMA, Nantes),
Nelle VAROQUAUX (Mines Paristech, Paris)
Benjamin VIALLE (CNIL, Paris)

■ **RÉSUMÉ** • L'évaluation formative pose des enjeux tant pédagogiques que logistiques. Elle nécessite en effet des dispositifs techniques efficaces, permettant aux enseignants et aux étudiants, d'accéder aux travaux rendus, de les corriger et de proposer un retour adapté. Elle est d'autant plus difficile à mettre en œuvre lorsque les volumes de travaux à corriger sont importants et les équipes pédagogiques hétérogènes. Pour répondre à ces besoins, nous présentons ici le logiciel libre MarkUs que nous co-développons avec l'Université de Toronto. Application web permettant de recueillir, annoter et évaluer quantitativement et qualitativement les productions des apprenants, MarkUs contribue à la rationalisation du processus de correction. Nous en détaillons l'impact dans le cadre des enseignements d'informatique menés à l'École Centrale de Nantes. Nous illustrons ainsi combien un effort pédagogique peut être soutenu par le déploiement d'un outil TICE lorsque ce dernier est en adéquation avec les attentes des enseignants.

■ **MOTS CLÉS** • évaluation formative, évaluation critériée, logiciel libre, correction en volumes, annotation, application web.

■ **ABSTRACT** • *Formative assessment raises both pedagogical and logistical outcomes. It requires effective tools allowing teachers and students to get access to submitted works, evaluate them and offer a suitable feedback. The complexity of this whole process increases with the growing number of works to assess or the heterogeneity of the team of teachers. In this paper, we introduce a new tool to address these issues: MarkUs is a free software that we are co-developing with the University of Toronto. It is a web application for collecting, annotating and quantitatively and qualitatively assessing the learners' productions. MarkUs is a major help to improve the assessment workflow. We detail its impact in the context of Computer Science courses at Ecole Centrale de Nantes. We thus illustrate how a pedagogical effort can be supported by the deployment of an ICT tool in line with teachers' expectations.*

■ **KEYWORDS** • *formative assessment, criterion-referenced assessment, free software, annotation, web application*

Sébastien GEORGE, Pierre JACOBONI, Christophe DESPRÉS, Françoise LE CALVEZ, Style pour la revue STICEF, *Revue STICEF*, en ligne sur www.sticef.org

1

1. Contexte et enjeux

1.1. La problématique de l'évaluation formative

Intégrée au déroulement d'un enseignement, l'évaluation formative vise à « *informer l'apprenant puis l'enseignant sur le degré d'atteinte des objectifs* » (Raynal et Rieunier, 1997). De nombreux articles et rapports ont mis l'accent sur ses avantages. Comme d'autres organismes avant elle (notamment l'OCDE en 2008 (OCDE-CERI, 2008)), la mission parlementaire Fourgous (Mission Fourgous, 2012) rappelle les atouts de l'évaluation formative pour aider à faire progresser les élèves : la possibilité pour l'enseignant d'ajuster son enseignement aux besoins des élèves, le fait d'associer activement les élèves au processus d'enseignement et d'apprentissage, ou encore l'opportunité donnée aux élèves d'élaborer des stratégies afin d'apprendre à apprendre. Les auteurs de (Black et Wiliam, 1998) concluent justement que l'évaluation formative améliore l'apprentissage, ajoutant que ces progrès figurent parmi les plus importants dont il ait jamais été fait état pour des interventions pédagogiques. La pertinence de l'évaluation formative dépend notamment de la capacité des étudiants et des professeurs à en prendre en compte les résultats pour ajuster le processus d'apprentissage. Ce qui implique non seulement un retour régulier des enseignants vers les apprenants pour les aider dans leur démarche, mais aussi un ajustement des activités pédagogiques aux compétences et connaissances développées par les élèves.

Dans (Closset et Delforge, 2011), les auteurs décrivent comment l'évaluation formative est intégrée aux dispositifs pédagogiques dans différents établissements. S'ils reconnaissent qu'il est difficile d'évaluer rigoureusement l'impact d'une telle démarche, leurs observations établissent une « *forte relation entre la moyenne aux tests formatifs et la réussite de l'année* ». Pour les auteurs, cette évaluation fournit aux étudiants un « *retour métacognitif sur leurs pratiques* », enrichissant dès lors qu'il est conjoint à un dispositif d'accompagnement. Ce dernier peut prendre plusieurs formes : mise en place d'une séance de correction dédiée, accueil des apprenants par l'équipe pédagogique pour répondre à leurs questions, mise en place d'un système de tutorat, ... La qualité d'une telle évaluation formative passe dès lors par un dispositif logistique adapté. L'augmentation des promotions d'élèves, l'hétérogénéité des équipes enseignantes (partagées entre permanents, attachés temporaires, doctorants et vacataires), la mobilité des enseignants-chercheurs (qui ne peuvent décemment pas emporter les copies des étudiants lors de leurs déplacements pour réunions ou colloques) ou encore la densification des séquences pédagogiques (compactées sur des périodes de temps réduites) sont autant de freins à une évaluation formative efficace. Pour y faire face, l'Université de Toronto a initié, dès 2006, un logiciel libre dédié : MarkUs (MarkUs). L'École Centrale de Nantes contribue à son développement depuis 2009 (Ma-

gnin et al., 2012), et le déploie dans l'ensemble des enseignements d'informatique depuis 2010.

1.2. Contribution

MarkUs est une application web destinée à simplifier la tâche de correction des schémas, des rapports et du code rendu par les étudiants dans le cadre de TP ou projets (initialement d'informatique, mais il s'étend aujourd'hui à d'autres domaines). Ce programme permet non seulement de réunir sur une même plate-forme le code remis par les élèves, mais aussi de pratiquer une évaluation critériée, accompagnée (ou non) d'une note selon une liste de facteurs définis par le responsable d'enseignement. Ainsi, les étudiants n'ont plus à imprimer leur code ou à l'envoyer par courriel aux enseignants : tout est géré via MarkUs, du dépôt des travaux des élèves à la correction ligne par ligne de leurs documents par les enseignants jusqu'au dévoilement des résultats. Qui plus est, cette solution s'inscrit dans l'effort général de développement durable : en offrant une manière efficace pour annoter et parcourir des documents dématérialisés, elle évite l'impression de documents temporaires de travail.

Dans cet article, nous détaillons l'apport de MarkUs dans les cours d'informatique de tronc commun proposés à l'École Centrale de Nantes, où le logiciel sert à suivre les travaux remis par des promotions de près de 400 étudiants. Dans un premier temps, nous décrivons le contexte de cette expérimentation, et les dispositifs qui ont été successivement adoptés pour assurer les processus d'évaluation formative. Nous présentons ensuite MarkUs et proposons une analyse critique de son utilisation. Nous expliquons en quoi la mise en œuvre de MarkUs a permis de soutenir les efforts des équipes pédagogiques. Enfin, nous dressons plusieurs perspectives pour un tel outil dans le but de renforcer son adéquation au cadre de l'évaluation formative.

2. Description du cadre, de la scénarisation des enseignements visés et des limites des dispositifs précédents

2.1. L'enseignement de l'informatique à l'École Centrale de Nantes

L'informatique fait partie des enseignements que reçoivent tous les élèves-ingénieurs à l'École Centrale de Nantes, notamment *via* des enseignements généraux d'algorithmique et de programmation pour l'ingénieur. L'informatique représente pour les deux années de tronc commun 64 heures d'enseignement, concernant environ 360 élèves par promotion répartis en douze groupes de travaux dirigés. Une quarantaine d'étudiants de troisième année choisit par ailleurs de se spécialiser en informatique afin d'approfondir leurs connaissances avec 200 heures de cours supplémentaires. En outre, chaque matière est enseignée par un ou deux enseignants de cours magistraux, des chargés de TP (travaux pratiques) et des chargés de TD (travaux dirigés). L'importance des promotions implique qu'il y ait un grand nombre de vacataires d'origines

diverses (doctorants, post-doctorants, extérieurs, enseignants d'autres départements). A titre d'exemple, l'enseignement de première année d'algorithmique et programmation (algorithmes de base et programmation en C), constitué de 32 heures de cours dont 16 heures de TP, nécessite sept permanents, neuf doctorants vacataires et deux extérieurs. Le responsable d'enseignement fait aussi face à un renouvellement du personnel important (entre cinq et sept nouvelles personnes participent à l'enseignement chaque année), ce qui entraîne des difficultés pour assurer la continuité de l'enseignement. S'ajoute à cela la nécessité d'une correction des travaux en moins de deux semaines, pour un suivi efficace des étudiants entre deux séances de TP.

2.2. Les précédents dispositifs déployés

Historiquement, le code et un rapport étaient imprimés, puis remis à l'enseignant à l'issue du TP ou déposés dans son casier. Ce dispositif présente de nombreux inconvénients, parmi lesquels un nombre élevé d'impressions, un exemplaire corrigé unique ne pouvant être partagé entre les différents élèves d'un groupe, une correction difficile pour un enseignant physiquement absent de l'école, et un suivi des retards peu efficace. Une autre méthode consistait à remettre par courriel à l'issue du TP le rapport et le code en format électronique dans une archive compressée. Ce procédé entraînait des complications techniques : archives mal formées, difficulté à annoter le code ligne par ligne, erreurs d'envoi dans les destinataires, boîtes mails surchargées. Par ailleurs, les élèves profitaient du système pour rendre leur travail en retard.

Les dispositifs précédemment décrits ne permettaient pas au responsable de matière d'avoir une vue globale du processus de correction : oubli de la part de certains enseignants de corriger un TP, disparité des notes selon les correcteurs, gestion des retards variable. Cela causait des sentiments d'injustice dans l'esprit des élèves qui avaient la sensation de ne pas être tous évalués selon les mêmes critères.

2.3. Le nouveau dispositif : MarkUs

2.3.1. Fonctionnalités

Initié en 2005 sous le nom d'OLM (Online Marking Tool) à l'Université de Toronto, MarkUs est une application web destinée à simplifier la tâche de correction du code rendu par les étudiants en informatique. Les responsables d'enseignement et les étudiants peuvent former des groupes de travail, et travailler sur des projets en utilisant un système de gestion de version (en l'occurrence SVN (Subversion)) par ligne de commande, ou via l'interface web. Celle-ci permet d'ajouter, de remplacer, ou de supprimer des fichiers très simplement. Le responsable de matière attribue ensuite des correcteurs aux groupes d'étudiants. Les dates butoir sont gérées directement au sein de l'application, qui permet de mettre en place un système de pénalité.

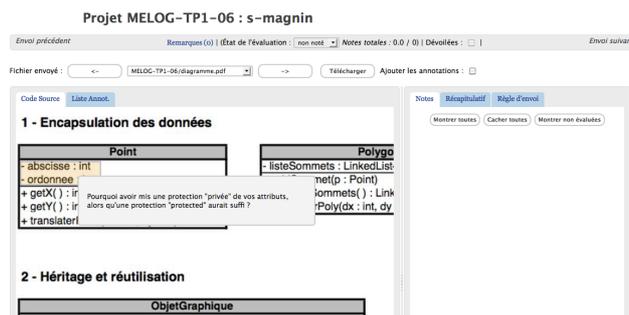


Figure 1 : Écran de correction des travaux d'un élève par un enseignant directement via un navigateur web. Ici, consultation et annotation d'un document PDF

Les correcteurs annotent le code, les images et les PDF déposés en ligne par les étudiants (cf. Figure 1). Ils bénéficient de la coloration syntaxique suivant le langage utilisé. MarkUs permet de dématérialiser la correction : les correcteurs accèdent aux devoirs depuis un navigateur web.

Comme le montre la Figure 2, l'évaluation peut se faire via une liste de critères et un barème établis par le responsable de matière.

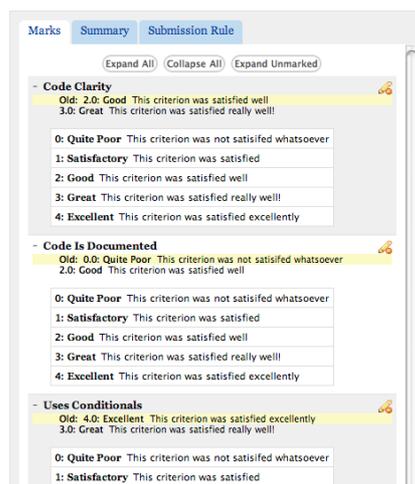


Figure 2 : Définition, par le responsable d'enseignement, d'une liste de critères d'évaluation

De plus, lorsque MarkUs vise à évaluer des travaux informatiques, il permet de réaliser une exécution automatique de tests unitaires. Cela signifie que le logiciel peut indiquer si le code remis par les étudiants répond correctement ou non à un ensemble

de tests définis par le responsable de matière. Cette fonctionnalité nouvelle n'est, pour le moment, pleinement utilisable que pour les enseignements portant sur le langage JAVA. Le responsable de matière peut visualiser facilement les travaux corrigés et ceux qui ne le sont pas encore. Une fois tous les travaux corrigés, le responsable de matière choisit d'afficher les résultats (notes et commentaires) aux étudiants.

2.3.2. Communauté et développement

Chaque semestre, une douzaine d'étudiants co-encadrés par des enseignants et des mentors techniques de l'Université de Toronto et de l'École Centrale de Nantes participe à l'implémentation de nouvelles fonctionnalités. Depuis 2008, plus de 45 étudiants ont ainsi pu participer au développement de MarkUs, dans le cadre de projets ou de stage. Une fois qu'ils ont acquis une bonne connaissance du code, ils peuvent devenir à leur tour mentor technique des nouveaux venus.

Courant septembre 2010, le code source de MarkUs est passé d'un hébergement privé à l'Université de Toronto à un hébergement public sur Github (Code MarkUs), permettant ainsi à de nouveaux développeurs de contribuer à l'amélioration du code de MarkUs. Bénéficiant d'une liste de diffusion publique, d'un wiki, de deux canaux IRC - l'un francophone et l'autre anglophone - et d'un blog, la communauté attache une attention particulière à une bonne communication ; ceci facilite les discussions afin d'assurer une haute qualité de code, et favorise l'intégration rapide des nouveaux développeurs et utilisateurs.

3. Analyse de l'impact de MarkUs sur l'enseignement de l'informatique

Dans cette section, nous analysons l'impact de MarkUs sur notre dispositif d'évaluation des travaux des étudiants en nous basant sur des critères aussi bien quantitatifs que qualitatifs.

3.1. Critères quantitatifs

En première année du cursus ingénieur (niveau Licence 3), MarkUs est utilisé dans le cadre du cours algorithmique et programmation qui comporte cinq travaux pratiques dont quatre donnent lieu à un compte rendu par les étudiants. L'évaluation étant faite par ailleurs, les comptes-rendus ne sont pas notés pour mettre en avant la notion d'entraînement pendant les travaux pratiques. L'équipe pédagogique, constituée de dix-huit personnes, utilise MarkUs depuis deux ans avec des promotions d'environ 360 élèves répartis en douze groupes. La première année a été consacrée à la mise en place du dispositif avec les enseignants et la deuxième année a permis le déploiement de MarkUs sur l'ensemble des groupes. Moins de 10% des travaux sont rendus en retard et environ 90% des compte-rendus sont corrigés en deux semaines, ce qui constitue une amélioration très notable par rapport à la situation précédente.

En seconde année du cursus ingénieur (niveau M1), MarkUs est utilisé dans le cadre d'un cours de tronc commun qui comporte trois TP systématiquement corrigés et remis aux élèves dans un délai inférieur à deux semaines. L'équipe pédagogique est composée de dix personnes et les promotions concernées sont de l'ordre de 360 élèves répartis en douze groupes. MarkUs est utilisé dans cet enseignement depuis l'année universitaire 2009-2010. Après une première année d'expérimentation sur la base du volontariat des chargés de TP, il est maintenant utilisé systématiquement. Un sondage effectué à l'issue de l'année scolaire 2010-2011 faisait état d'une diminution du temps de correction de 15 à 50% suivant les enseignants. De la même manière, la proportion d'étudiants rendant leurs travaux en retard est passée de 15-20% à seulement 5-10%, la machine ayant la réputation d'être plus « inflexible » que l'enseignant.

Le responsable de matière peut également vérifier le délai de correction, voire la cohérence des corrections rendues aux étudiants et avertir son équipe pédagogique des éventuelles mesures à prendre. Les TP de cette matière ne sont pas notés, ce choix a été effectué pour responsabiliser les étudiants qui ne cherchent ainsi plus simplement la bonne note (éventuellement en fraudant) mais se focalisent plus sur la compréhension du problème posé. Enfin, l'estimation effectuée de l'économie de papier imprimé est d'environ 5000 feuilles par matière et par an.

3.2. Soutien de l'évaluation formative

Dans le cours de première année, l'importance d'une évaluation formative pour les étudiants débutants est déterminante pour l'acquisition de bases saines en informatique. En effet, les étudiants ayant des cursus d'origines très variés (classes préparatoires, université ou autres), leur niveau en informatique est très hétérogène. Les groupes de TD étant formés sur la base des compétences de langues pour faciliter l'organisation des enseignements, cette hétérogénéité se retrouve au sein d'un même groupe de TD. L'évaluation formative est donc particulièrement adaptée à la situation à laquelle est confrontée l'équipe pédagogique. MarkUs facilite et accélère l'évaluation par la possibilité d'avoir une liste d'annotations prédéfinies qui permettent en plus une certaine normalisation des critères d'évaluation et permettent de vérifier facilement l'acquisition de certaines compétences (par l'absence de certaines erreurs typiques). Les annotations prédéfinies (dont on voit un exemple sur la Figure 3) ne sont bien sûr pas exhaustives, et elles peuvent être complétées des annotations personnelles de chaque enseignant.

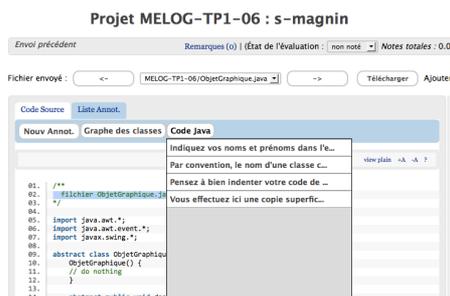


Figure 3 : Accès par un correcteur aux annotations prédéfinies par le responsable de matière

Du côté étudiant, MarkUs fournit également un soutien en conservant la mémoire des travaux déposés : les étudiants conservent tous les corrections (contrairement aux corrections sur papier où seul un membre du binôme conservait cette copie) et peut consulter à tout moment de l'année à la fois les fichiers rendus et les éléments de correction fournis par le chargé de TD/TP.

De par les contraintes de temps et les occupations variées des enseignants-chercheurs, il est indispensable qu'ils puissent corriger leur TP depuis leur lieu de travail, comme de leur domicile mais aussi lors de leurs déplacements, indépendamment de la plate-forme matérielle et logicielle qu'ils utilisent. MarkUs, en tant qu'application Web répond parfaitement à cette demande, elle peut s'utiliser depuis n'importe quel matériel connecté à internet, ordinateur portable ou même tablette. Enfin, il fournit une vue globale au responsable de l'équipe pédagogique qui peut détecter ainsi facilement des disparités entre les différents groupes (délais de rendu des TP ou des corrections, notes, types d'annotations...).

3.3. Accompagnement de l'équipe enseignante

MarkUs se connecte facilement avec le système d'information de l'établissement : il récupère ainsi les listes des inscrits à un cours et les différents groupes qui en découlent. L'identification des étudiants est assurée avec le référentiel LDAP de l'établissement. Il devient ainsi un outil standard, facile à utiliser et constitue l'outil unique de dépôt – consultation et correction des travaux des étudiants (en pratique il est également utilisé dans des matières à plus faible effectif au niveau M2).

Trois supports d'utilisation de MarkUs ont été réalisés pour faciliter la prise en main de l'application par les étudiants et l'équipe pédagogique : ils couvrent respectivement le rôle de l'étudiant, d'un chargé de TP et du responsable de la matière.

3.4. Limites

L'utilisation de MarkUs en contexte opérationnel a certes permis d'améliorer le suivi et la correction des travaux pratiques, mais aussi de révéler quelques limites. Concernant l'évaluation formative, il est possible d'établir une liste de critères communs à tous les projets auxquels participe un étudiant, et donc de constater sa progression tout au long du module. Cependant, MarkUs ne peut pas être considéré comme un outil d'évaluation individuelle, car il ne permet pas encore d'obtenir une carte automatique des progrès de l'étudiant. En intégrant une liste de compétences à acquérir à la fin d'un enseignement, nous pourrions établir un suivi de l'acquisition de ces compétences projet par projet. Par ailleurs, cette validation d'acquisition de compétences est toujours limitée par le fait que les travaux pratiques se font en binôme, et ne permet pas de savoir lequel des étudiants a réellement acquis les compétences en question. Cette limitation n'est pas liée directement au système MarkUs, mais ce dernier pourrait malgré tout aider les enseignants à faire une évaluation individuelle.

Concernant l'authenticité des travaux rendus, les étudiants sont malheureusement très touchés par le plagiat, et il est évident qu'un système de détection automatique de plagiat serait une aide précieuse pour les enseignants. Un module de détection de plagiat est en cours de développement pour répondre à ce besoin. Cependant, la simple détection de plagiat et les mesures qui suivent en cas de fraude avérée (invalidation des TP, voire du module complet selon les cas) n'ont que peu d'efficacité si elles ne sont pas accompagnées d'un volet pédagogique sur la nécessité de rendre un travail authentique. La question est donc de savoir comment un système tel que MarkUs pourrait nous aider à associer un aspect pédagogique à la détection de plagiat.

Ces limites constituent une partie du cahier des charges pour les évolutions du système. Étant partie prenante dans le développement de ce logiciel libre qu'est MarkUs, nous avons la chance de pouvoir intervenir directement sur son amélioration incrémentale. Les bénéfices d'un tel mode de fonctionnement ont été décrits dans (Magnin et Moreau, 2011).

4. Synthèse et perspectives

MarkUs est utilisé à l'Université de Toronto, à l'Université de Waterloo et à l'École Centrale de Nantes. Dans cette dernière, nous avons déployé MarkUs pour des enseignements de tronc commun, c'est-à-dire impactant à chaque fois près de 400 étudiants. Les élèves se sont avérés beaucoup plus respectueux des échéances qu'ils ne l'étaient auparavant, ont porté plus d'intérêt aux annotations laissées sur la plateforme, et ont apprécié la vitesse accrue de correction de la part de leurs chargés de TD/TP. Les enseignants ont, pour leur part, tiré parti des annotations et critères partagés au niveau global de l'équipe pédagogique. Ils ont été sensibles à la disponibilité permanente de la plate-forme et à la possibilité de corriger depuis tout lieu, qu'il s'agisse de leur bureau, de leur maison ou d'une salle située à l'autre bout du monde.

Surtout, MarkUs a permis d'unifier les critères de correction d'un groupe à l'autre et de fournir une évaluation formative beaucoup plus efficace et pertinente. MarkUs suscite désormais l'intérêt de différents établissements européens, et nous sommes heureux d'accueillir, au sein de l'équipe de développement, de nouveaux contributeurs. Si les expériences d'utilisation de MarkUs se sont jusqu'à maintenant concentrées sur les enseignements d'informatique, son ouverture et sa généralisation au support de fichiers PDF de grande taille (une fonctionnalité développée par deux élèves contributeurs entre janvier et mars 2012) vont permettre de l'utiliser dans tout type de matières. Plusieurs de nos collègues ont déjà manifesté leur intérêt, notamment pour la gestion des rapports de stages ou de projets industriels.

Bibliographie et webographie

(Black et Wiliam, 1998)

BLACK P, WILAM D. (1998). *Assessment and Classroom Learning, Assessment in Education: Principles, Policy and Practice*, CARFAX, Oxfordshire, Vol. 5, No. 1. 1998.

(Closset et Delforge, 2011)

CLOSSET J.L., DELFORGE M. (2011). Comment tirer parti d'une évaluation formative ? In Ph. Parmentier (Dir.), *Recherches et actions en faveur de la réussite en première année universitaire. Vingt ans de collaboration dans la Commission « Réussite » du Conseil interuniversitaire de la Communauté française de Belgique*. Bruxelles : CIUF. 2011.

(Magnin et al., 2012)

MAGNIN M., MOREAU G., VAROQUAUX N., VIALLE B., REID K., CONLEY M., GEHWOLF S. (2012). MarkUs, an open-source web application to annotate student papers online. *Proceedings of the ASME 11th Biennial Conference On Engineering Systems Design And Analysis (ESDA)*. Nantes, France. 2012.

(Magnin et Moreau, 2011)

MAGNIN M., MOREAU G. (2011). La professionnalisation des étudiants en informatique par les projets libres. Colloque *Questions de pédagogie dans l'enseignement supérieur (QPES 2011)*, 6 pages, Angers, Juin 2011.

(Mission Fourgous, 2012)

Mission Fourgous. Apprendre autrement à l'ère du numérique. *Rapport parlementaire*. France. 2012.

(OCDE-CERI, 2008)

OCDE-CERI. Évaluer l'apprentissage : L'évaluation formative. *Rapport de l'OCDE*. 2008.

(Raynal et Rieunier, 1997)

RAYNAL F., RIEUNIER A. (1997). F. Raynal et A. Rieunier. *Pédagogie : dictionnaire des concepts clés*. Paris : ESF. 1997.

(MarkUs)

Site officiel du projet MarkUs : <http://markusproject.org> (consulté le 18 octobre 2012)

(Code MarkUs)

Dépôt collaboratif du projet MarkUs, contenant l'historique du projet et la documentation : <https://github.com/MarkUsProject/Markus> (consulté le 18 octobre 2012)

(Subversion)

Site officiel du projet Subversion, hébergé par la fondation Apache : <http://subversion.apache.org/> (consulté le 18 octobre 2012)